

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q 1/34, C12N 11/00, C07K 7/06	A1	(11) International Publication Number: WO 98/20156 (43) International Publication Date: 14 May 1998 (14.05.98)
(21) International Application Number: PCT/CA97/00825 (22) International Filing Date: 3 November 1997 (03.11.97) (30) Priority Data: 60/030,408 4 November 1996 (04.11.96) US (71) Applicant (for all designated States except US): MERCK FROSST CANADA INC. [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). (72) Inventors; and (75) Inventors/Applicants (for US only): DESMARAIS, Sylvie [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). FRIESEN, Richard [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). GRESSER, Michael [US/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). KENNEDY, Brian [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). NICHOLSON, Don [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). RAMACHANDRAN, Chidambaran [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). SKOREY, Kathryn [CA/CA]; 16711 Trans Canada Highway, Kirkland, Quebec H9H 3L1 (CA). FORD-HUTCHINSON,	(74) Agent: MURPHY, Kevin, P.; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College, Montreal, Quebec H3A 2Y3 (CA). (81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: HYDROLASE BINDING ASSAY (57) Abstract <p>Disclosed is a binding assay for proteases and phosphatases, which contain cysteine in their binding sites or as a necessary structural component for enzymatic binding. The sulfhydryl group of cysteine is the nucleophilic group in the enzyme's mechanistic proteolytic and hydrolytic properties. The assay can be used to determine the ability of new, unknown ligands and mixtures of compounds to competitively bind with the enzyme versus a known binding agent for the enzyme, e.g., a known enzyme inhibitor. By the use of a mutant form of the natural or native wild-type enzyme, in which serine, or another amino acid, e.g., alanine, replaces cysteine, the problem of interference from extraneous oxidizing and alkylating agents in the assay procedure is overcome. The interference arises because of oxidation or alkylation of the sulfhydryl, -SH (or -S⁻), in the cysteine, which then adversely affects the binding ability of the enzyme. Specifically disclosed is an assay for tyrosine phosphatases and cysteine proteases, including capsases and cathepsins, e.g., Cathepsin K(O2), utilizing scintillation proximity assay (SPA) technology. The assay has important applications in the discovery of compounds for the treatment and study of, for example, diabetes, immunosuppression, cancer, Alzheimer's disease and osteoporosis. The novel feature of the use of a mutant enzyme can be extended to its use in a wide variety of conventional colorimetric, photometric, spectrophotometric, radioimmunoassay and ligand-binding competitive assays.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE OF THE INVENTION
HYDROLASE BINDING ASSAY

5

FIELD OF THE INVENTION

This invention relates to the use of mutant phosphatase
10 and protease enzymes in a competitive binding assay. Specific
examples are the enzymes, tyrosine phosphatase and cysteine
protease, e.g. Cathepsin K, and the assay specifically described is a
scintillation proximity assay using a radioactive inhibitor to induce
scintillation.

15 BACKGROUND OF THE INVENTION

The use of the scintillation proximity assay (SPA) to
study enzyme binding and interactions is a new type of
radioimmunoassay and is well known in the art. The advantage of
SPA technology over more conventional radioimmunoassay or
20 ligand-binding assays, is that it eliminates the need to separate
unbound ligand from bound ligand prior to ligand measurement. See
for example, *Nature*, Vol, 341, pp. 167-178 entitled "Scintillation
Proximity Assay " by N. Bosworth and P. Towers, *Anal. Biochem.*
Vol. 217, pp. 139-147 (1994) entitled "Biotinylated and Cysteine-
25 Modified Peptides as Useful Reagents For Studying the Inhibition of
Cathepsin G" by A.M. Brown, et al., *Anal. Biochem.* Vol. 223, pp. 259-
265 (1994) entitled "Direct Measurement of the Binding of RAS to
Neurofibromin Using Scintillation Proximity Assay" by R. H.
Skinner *et al.* and *Anal. Biochem.* Vol. 230, pp. 101-107(1995) entitled
30 "Scintillation Proximity Assay to Measure Binding of Soluble

Fibronectin to Antibody-Captured $\alpha_5\beta_1$ Integrin" by J. A. Pachter *et al.*

The basic principle of the assay lies in the use of a solid support containing a scintillation agent, wherein a target enzyme is attached to the support through, e.g., a second enzyme-antienzyme linkage. A known tritiated or I^{125} iodinated binding agent, i.e., radioligand inhibitor ligand for the target enzyme is utilized as a control, which when bound to the active site in the target enzyme, is in close proximity to the scintillation agent to induce a scintillation signal, e.g., photon emission, which can be measured by conventional scintillation/radiographic techniques. The unbound tritiated (hot) ligand is too far removed from the scintillation agent to cause an interfering measurable scintillation signal and therefore does not need to be separated, e.g., filtration, as in conventional ligand-binding assays.

The binding of an unknown or potential new ligand (cold, being non-radioactive) can then be determined in a competitive assay versus the known radioligand, by measuring the resulting change in the scintillation signal which will significantly decrease when the unknown ligand also possesses good binding properties.

However, a problem arises when utilizing a target enzyme containing a cysteine group, having a free thiol linkage, -SH, (or present as $-S^-$) which is in the active site region or is closely associated with the active site and is important for enzyme-ligand binding. If the unknown ligand or mixture, e.g. natural product extracts, human body fluids, cellular fluids, etc. contain reagents which can alkylate, oxidize or chemically interfere with the cysteine thiol group such that normal enzyme-ligand binding is disrupted, then false readings will occur in the assay.

What is needed in the art is a method to circumvent and avoid the problem of cysteine interference in the scintillation proximity assay (SPA) procedure in enzyme binding studies.

SUMMARY OF THE INVENTION

We have discovered that by substituting serine for cysteine in a target enzyme, where the cysteine plays an active role in the wild-type enzyme-natural ligand binding process, usually as the catalytic nucleophile in the active binding site, a mutant is formed which can be successfully employed in a scintillation proximity assay without any active site cysteine interference.

This discovery can be utilized for any enzyme which contains cysteine groups important or essential for binding and/or catalytic activity as proteases or hydrolases and includes phosphatases, e.g., tyrosine phosphatases and proteases, e.g. cysteine proteases, including the cathepsins, i.e., Cathepsin K (O2) and the capsases.

Further, use of the mutant enzyme is not limited to the scintillation proximity assay, but can be used in a wide variety of known assays including colorimetric, spectrophotometric, ligand-binding assays, radioimmunoassays and the like.

We have furthermore discovered a new method of amplifying the effect of a binding agent ligand, e.g., radioactive inhibitor, useful in the assay by replacing two or more phosphotyrosine residues with 4-phosphono(difluoromethyl) phenylalanine (F₂Pmp) moieties. The resulting inhibitor exhibits a greater and more hydrolytically stable binding affinity for the target enzyme and a stronger scintillation signal.

By this invention there is provided a process for determining the binding ability of a ligand to a cysteine-containing wild-type enzyme comprising the steps of:

- (a) contacting a complex with the ligand, the complex comprising a mutant form of the wild-type enzyme, in which cysteine, at the active site, is replaced with serine, in the presence of a known binding agent for the mutant enzyme, wherein the binding agent is capable of binding with the mutant enzyme to produce a measurable signal.

Further provided is a process for determining the binding ability of a ligand, preferably a non-radioactive (cold) ligand, to an active site cysteine-containing wild-type tyrosine phosphatase comprising the steps of:

- 5 (a) contacting a complex with the ligand, the complex comprising a mutant form of the wild-type enzyme, the mutant enzyme being PTP1B, containing the same amino acid sequence 1-320 as the wild type enzyme, except at position 215, in which cysteine is
- 10 replaced with serine in the mutant enzyme, in the presence of a known radioligand binding agent for the mutant enzyme, wherein the binding agent is capable of binding with the mutant enzyme to produce a measurable beta radiation-induced
- 15 scintillation signal.

Also provided is a new class of peptide binding agents selected from the group consisting of:

- 20 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide (BzN-EJJ-CONH₂), where E is glutamic acid and J is 4-phosphono(difluoro-methyl)-L-phenylalanyl;
- N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 25 N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 30 L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide; and
- 35 L-Isoleucinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide; and their tritiated and I¹²⁵ iodinated derivatives.

Further provided is a novel tritiated peptide, tritiated BzN-EJJ-CONH₂, being N-(3,5-Ditritio)benzoyl-L-glutamyl-[4-phosphono-(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide, wherein E as used herein is
5 glutamic acid and J, as used herein, is the (F₂Pmp) moiety, (4-phosphono(difluoromethyl)-phenylalanyl).

Furthermore there is provided a process for increasing the binding affinity of a ligand for a tyrosine phosphatase or cysteine protease comprising introducing into the ligand two or more 4-
10 phosphono(difluoromethyl)-phenylalanine groups; also provided is the resulting disubstituted ligand.

In addition there is provided a complex comprised of:

- (a) a mutant form of a wild-type enzyme, in which cysteine, necessary for activity in the active site, is
15 replaced with serine and is attached to:
- (b) a solid support.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 illustrates the main elements of the invention
20 including the scintillation agent 1, the supporting (fluomicrosphere) bead 5, the surface binding Protein A 10, the linking anti-GST enzyme 15, the fused enzyme construct 20, the GST enzyme 25, the mutant enzyme 30, the tritiated peptide inhibitor 35, the beta radiation emission 40 from the radioactive peptide inhibitor 35 and
25 the emitted light 45 from the induced scintillation.

FIGURE 2 (A and B) illustrates the DNA and amino acid sequences for PTP1B tyrosine phosphatase enzyme, truncated to amino acid positions 1-320. (Active site cysteine at position 215 is in
30 bold and underlined).

FIGURE 3 (A, B and C) illustrates the DNA and amino acid sequences for Cathepsin K. The upper nucleotide sequence represents the cathepsin K cDNA sequence which encodes the
35 cathepsin K preproenzyme (indicated by the corresponding three letter amino acid codes). Numbering indicates the cDNA nucleotide

position. The underlined amino acid is the active site Cys¹³⁹ residue that was mutated to either Ser or Ala.

FIGURE 4 (A and B) illustrates the DNA and amino acid sequences for the capsase, apopain. The upper nucleotide sequence represents the apopain (CPP32) cDNA sequence which encodes the apopain proenzyme (indicated by the corresponding three letter amino acid codes). Numbering indicates the cDNA nucleotide position. The underlined amino acid is the active site Cys¹⁶³ residue that was mutated to Ser.

DETAILED DESCRIPTION OF THE INVENTION

The theory underlying the main embodiment of the invention can be readily seen and understood by reference to

FIGURE 1.

Scintillation agent 1 is incorporated into small (yttrium silicate or PVT fluomicro-spheres, AMERSHAM) beads 5 that contain on their surface immunosorbent protein A 10. The protein A coated bead 5 binds the GST fused enzyme construct 20, containing GST enzyme 25 and PTP1B mutant enzyme 30, via anti-GST enzyme antibody 15. When the radioactive e.g., tritiated, peptide 35 is bound to the mutant phosphatase enzyme 30, it is in close enough proximity to the bead 5 for its beta emission 40 (or Auger electron emission in the case of I¹²⁵) to stimulate the scintillation agent 1 to emit light (photon emission) 45. This light 45 is measured as counts in a beta plate counter. When the tritiated peptide 35 is unbound it is too distant from the scintillation agent 1 and the energy is dissipated before reaching the bead 5, resulting in low measured counts. Non-radioactive ligands which compete with the tritiated peptide 35 for the same binding site on the mutant phosphatase enzyme 30 will remove and/or replace the tritiated peptide 35 from the mutant enzyme 30 resulting in lower counts from the uncompeted peptide control. By varying the concentration of the unknown ligand and measuring the resulting lower counts, the inhibition at 50%(IC₅₀) for ligand binding to the mutant enzyme 30 can be obtained. This then is a measure of

the binding ability of the ligand to the mutant enzyme and the wild-type enzyme.

The term "complex" as used herein refers to the assembly containing the mutant enzyme. In its simplest embodiment, the complex is a solid support with the mutant enzyme attached to the surface of the support. A linker can also be employed. As illustrated in FIGURE 1, the complex can further comprise a bead (fluopolymer), anti-enzyme GST/enzyme GST-mutant enzyme-PTP1 linking construct, immunosorbent protein A, and scintillation agent. In general, the complex requires a solid support (beads, immunoassay column of e.g., Al_2O_3 , or silica gel) to which the mutant enzyme can be anchored or tethered by attachment through a suitable linker, e.g., an immunosorbent (e.g., Protein A, Protein G, anti-mouse, anti-rabbit, anti-sheep) and a linking assembly, including an enzyme/anti-enzyme construct attached to the solid support.

The term "cysteine-containing wild-type enzyme", as used herein, includes all native or natural enzymes, e.g., phosphatases, cysteine proteases, which contain cysteine in the active site as the active nucleophile, or contain cysteine clearly associated with the active site that is important in binding activity.

The term "binding agent" as used herein includes all ligands (compounds) which are known to be able to bind with the wild-type enzyme and usually act as enzyme inhibitors. The binding agent carries a signal producing agent, e.g., radionuclide, to initiate the measurable signal. In the SPA assay the binding agent is a radioligand.

The term "measurable signal" as used herein includes any type of generated signal, e.g., radioactive, colorimetric, photometric, spectrophotometric, scintillation, which is produced when binding of the radioligand binding agent to the mutant enzyme.

The present invention assay further overcomes problems encountered in the past, where compounds were evaluated by their ability to affect the reaction rate of the enzyme in the phosphatase activity assay. However this did not give direct evidence that compounds were actually binding at the active site of the enzyme. The herein described invention binding assay using a substrate

analog can determine directly whether the mixtures of natural products can irreversibly modify the active site cysteine in the target enzyme resulting in inhibition of the enzymatic activity. To overcome inhibition by these contaminants in the phosphatase assay, a mutated
5 Cys(215) to Ser(215) form of the tyrosine phosphatase PTP1B was cloned and expressed resulting in a catalytically inactive enzyme. In general, replacement of cysteine by serine will lead to a catalytically inactive or substantially reduced activity mutant enzyme.

10 PTP1B is the first protein tyrosine phosphatase to be purified to near homogeneity (Tonks *et al.* *JBC* 263, 6731-6737 (1988)) and sequenced by Charbonneau *et al.* *PNAS* 85, 7182-7186 (1988). The sequence of the enzyme showed substantial homology to a duplicated domain of an abundant protein present in hematopoietic cells
15 variously referred to as LCA or CD45. This protein was shown to possess tyrosine phosphatase activity (Tonks *et al.* *Biochemistry* 27, 8695-8701 (1988)). Protein tyrosine phosphatases have been known to be sensitive to thiol oxidizing agents and alignment of the sequence of PTP1B with subsequently cloned *Drosophila* and mammalian
20 tyrosine phosphatases pointed to the conservation of a Cysteine residue ((M. Strueli *et al.* *Proc. Nat'l Acad USA*, Vol. 86, pp. 8698-7602 (1989)) which when mutated to Ser inactivated the catalytic activity of the enzymes. Guan *et al.* (1991) (*J.B.C.* Vol. 266, 17926-17030, 1991) cloned the rat homologue of PTP1B, expressed a truncated version of
25 the protein in bacteria, purified and showed the Cys at position 215 is the active site residue. Mutation of the Cys²¹⁵ to Ser²¹⁵ resulted in loss of catalytic activity. Human PTP1B was cloned by Chernoff *et al.* *Proc. Natl. Acad. Sci. USA* 87, 2735-2739 (1990).

Work leading up to the development of the substrate
30 analog BzN-EJJ-CONH₂ for PTP1B was published by T. Burke *et al.* *Biochem. Biophys. Res. Comm.* 205, pp. 129-134 (1994) with the synthesis of the hexamer peptide containing the phosphotyrosyl mimetic F₂Pmp. We have incorporated the (F₂Pmp) moiety (4-phosphono-(difluoromethyl)phenylalanyl) into various peptides that
35 led to the discovery of BzN-EJJ-CONH₂, (where E is glutamic acid and J as used herein is the F₂Pmp moiety) an active (5 nM) inhibitor

of PTP1B. This was subsequently tritiated giving the radioactive substrate analog required for the binding assay.

The mutated enzyme, as the truncated version, containing amino acids 1-320 (see FIGURE 2), has been demonstrated to bind the substrate analog Bz-NEJJ-CONH₂ with high affinity for the first time. The mutated enzyme is less sensitive to oxidizing agents than the wild-type enzyme and provides an opportunity to identify novel inhibitors for this family of enzymes. The use of a mutated enzyme to eliminate interfering contaminants during drug screening is not restricted to the tyrosine phosphatases and can be used for other enzyme binding assays as well.

Other binding assays exist in the art in which the basic principle of this invention can be utilized, namely, using a mutant enzyme in which an important and reactive cysteine important for activity can be modified to serine (or a less reactive amino acid) and render the enzyme more stable to cysteine modifying reagents, such as alkylating and oxidizing agents. These other ligand-binding assays include, for example, colorimetric and spectrophotometric assays, e.g. measurement of produced color or fluorescence, phosphorescence (e.g. ELISA, solid absorbant assays) and other radioimmunoassays in which short or long wave light radiation is produced, including ultraviolet and gamma radiation).

Further, the scintillation proximity assay can also be practiced without the fluopolymer support beads (AMERSHAM) as illustrated in FIGURE 1. For example, Scintistrips® are commercially available (Wallac Oy, Finland) and can also be employed as the scintillant-containing solid support for the mutant enzyme complex as well as other solid supports which are conventional in the art.

The invention assay described herein is applicable to a variety of cysteine-containing enzymes including protein phosphatases, proteases, lipases, hydrolases, and the like.

The cysteine to serine transformation in the target enzyme can readily be accomplished by analogous use of the molecular cloning technique for Cys²¹⁵ to Ser²¹⁵ described in the below-cited reference by M. Strueli *et al.*, for PTP1B and is hereby incorporated by reference for this particular purpose.

A particularly useful class of phosphatases is the tyrosine phosphatases since they are important in cell function. Examples of this class are: PTP1B, LCA, LAR, DLAR, DPTP(See Strueli et al., below). Ligands discovered by this assay using, for example, PTP1B can be useful, for example, in the treatment of diabetes and immunosuppression.

A useful species is PTP1B, described in *Proc. Nat'l Acad USA*, Vol. 86, pp. 8698-7602 by M. Strueli *et al.* and *Proc. Nat'l Acad Sci. USA*, Vol 87, pp. 2735-2739 by J. Chernoff *et al.*

Another useful class of enzymes is the proteases, including cysteine proteases (thiol proteases), cathepsins and capsases.

The cathepsin class of cysteine proteases is important since Cathepsin K (also termed Cathepsin O2, see *Biol. Chem. Hoppe-Seyler*, Vol. 376 pp. 379-384, June 1995 by D. Bromme *et al.*) is primarily expressed in human osteoclasts and therefore this invention assay is useful in the study and treatment of osteoporosis. See US Patent 5,501,969 (1996) to Human Genome Sciences for the sequence, cloning and isolation of Cathepsin K (O2). See also *J. Biol. Chem.* Vol. 271, No. 21, pp. 12511-12516 (1996) by F. Drake *et al.* and *Biol. Chem. Hoppe-Seyler*, Vol. 376, pp. 379-384(1985) by D. Bromme *et al.*, *supra*.

Examples of the cathepsins include Cathepsin B, Cathepsin G, Cathepsin J, Cathepsin K(O2), Cathesin L, Cathepsin M, Cathepsin S.

The capsase family of cysteine proteases are other examples where the SPA technology and the use of mutated enzymes can be used to determine the ability of unknown compounds and mixtures of compounds to compete with a radioactive inhibitor of the enzyme. An active site mutant of Human Apopain CPP32 (capsase-3) has been prepared. The active site thiol mutated enzymes are less sensitive to oxidizing agents and provide an opportunity to identify novel inhibitors for this family of enzymes.

Examples of the capsase family include: capsase-1(ICE), capsase-2 (ICH-1), capsase-3 (CPP32, human apopain, Yama), capsase-4(ICE_{rel}-11, TX, ICH-2), capsase-5(ICE_{rel}-111, TY), capsase-

6(Mch2), capsase-7(Mch3, ICE-LAP3, CMH-1), capsase-8(FLICE, MACH, Mch5), capsase-9 (ICE-LAP6, Mch6) and capsase-10(Mch4).

Substitution of the cysteine by serine (or by any other amino acid which lowers the activity to oxidizing and alkylating agents, e.g., alanine) does not alter the binding ability of the mutant enzyme to natural ligands. The degree of binding, i.e., binding constant, may be increased or decreased. The catalytic activity of the mutant enzyme will, however, be substantially decreased or even completely eliminated. Thus, natural and synthetic ligands which bind to the natural wild-type enzyme will also bind to the mutant enzyme.

Substitution by serine for cysteine also leads to the mutant enzyme which has the same qualitative binding ability as the natural enzyme but is significantly reduced in catalytically activity. Thus, this invention assay is actually measuring the true binding ability of the test ligand.

The test ligand described herein is a new ligand potentially useful in drug screening purposes and its mode of action is to generally function as an inhibitor for the enzyme.

The binding agent usually is a known ligand used as a control and is capable of binding to the natural wild-type enzyme and the mutant enzyme employed in the assay and is usually chosen as a known peptide inhibitor for the enzyme.

The binding agent also contains a known signal-producing agent to cause or induce the signal in the assay and can be an agent inducing e.g., phosphorescence or fluorescence (ELISA), color reaction or a scintillation signal.

In the instant embodiment, where the assay is a scintillation assay, the signal agent is a radionuclide, i.e., tritium, I^{125} , which induces the scintillant in the solid support to emit measurable light radiation, i.e., photon emission, which can be measured by using conventional scintillation and beta radiation counters.

We have also discovered that introducing two or more 4-phosphonodifluoromethyl phenylalanine (F₂Pmp) groups into a known binding agent greatly enhances the binding affinity of the

binding agent to the enzyme and improves its stability by rendering the resulting complex less susceptible to hydrolytic cleavage.

A method for introducing one F₂Pmp moiety into a ligand is known in the art and is described in detail in *Biochem.*

- 5 *Biophys. Res. Comm.* Vol. 204, pp. 129-134 (1994) hereby incorporated by reference for this particular purpose.

As a result of this technology we discovered a new class of ligands having extremely good binding affinity for PTP1B. These include:

- 10 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
- 15 L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
- 20 L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide, and
L-Isoleucinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide.
- 25 A useful ligand in the series is Bz-NEJJ-CONH₂, whose chemical name is: N-Benzoyl-L-glutamyl-[4-phosphono(difluoro-methyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenyl-alanineamide, and its tritiated form, N-(3,5-Ditritio)benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono-
- 30 (difluoromethyl)]-L-phenylalanineamide.

Synthesis of both cold and hot ligands is described in the Examples.

- The following Examples are illustrative of carrying out the invention and should not be construed as being limitations on the scope or spirit of the instant invention.
- 35

EXAMPLES

1. Preparation of PTP1B Truncate (Amino Acid Sequence from 1-320) and Fused GST-PTP1B Construct

- 5 An *E. coli* culture carrying a PET plasmid expressing the full length PTP1B protein was disclosed in J. Chernoff *et al. Proc Natl. Acad. Sci. USA* , 87, pp. 2735-2739, (1990). This was modified to a truncated PTP1B enzyme complex containing the active site with amino acids 1-320 inclusive, by the following procedure:
- 10 The full length human PTP-1B cDNA sequence (published in J. Chernoff et al., PNAS, USA, *supra*) cloned into a PET vector was obtained from Dr. Raymond Erickson (Harvard University). The PTP-1B cDNA sequence encoding amino acids 1-320 (Seq. ID No. 1) was amplified by PCR using the full length sequence
- 15 as template. The 5' primer used for the amplification included a Bam HI site at the 5' end and the 3' primer had an Eco RI site at the 3' end. The amplified fragment was cloned into pCR2 (Invitrogen) and sequenced to insure that no sequence errors had been introduced by Taq polymerase during the amplification. This sequence was
- 20 released from pCR2 by a Bam HI/Eco RI digest and the PTP-1B cDNA fragment ligated into the GST fusion vector pGEX-2T (Pharmacia) that had been digested with the same enzymes. The GST-PTP-1B fusion protein expressed in *E. Coli* has an active protein tyrosine phosphatase activity. This same 1-320 PTP-1B sequence (Seq. ID No.
- 25 1) was then cloned into the expression vector pFLAG-2, where FLAG is the octa-peptide AspTyrLysAspAspAspAspLys. This was done by releasing the PTP-1B sequence from the pGEX-2T vector by Nco I/Eco RI digest, filling in the ends of this fragment by Klenow and blunt-end ligating into the blunted Eco RI site of pFLAG2. Site-directed
- 30 mutagenesis was performed on pFLAG2-PTP-1B plasmid using the Chameleon (Stratagene) double-stranded mutagenesis kit from Stratagene, to replaced the active-site Cys-215 with serine. The mutagenesis was carried out essentially as described by the manufacturer and mutants identified by DNA sequencing. The
- 35 FLAG-PTP-1B Cys215Ser mutant (Seq. ID No. 7) was expressed, purified and found not to have any phosphatase activity. The GST-

PTP-1B Cys²¹⁵Ser mutant was made using the mutated Cys²¹⁵Ser sequence of PTP-1B already cloned into pFLAG2, as follows. The pFLAG2- PTP-1B Cys²¹⁵Ser plasmid (Seq. ID No. 7) was digested with Sal I (3' end of PTP-1B sequence), filled in using Klenow
5 polymerase (New England Biolabs), the enzymes were heat inactivated and the DNA redigested with Bgl II. The 500 bp 3' PTP-1B cDNA fragment which is released and contains the mutated active site was recovered. The pGEX-2T-PTP-1B plasmid was digested with Eco RI (3' end of PTP-1B sequence), filled in by Klenow,
10 phenol/chloroform extracted and ethanol precipitated. This DNA was then digested with Bgl II, producing two DNA fragments a 500 bp 3' PTP-1B cDNA fragment that contains the active site and a 5.5 Kb fragment containing the pGEX-2T vector plus the 5' end of PTP-1B. The 5.5 Kb pGEX-2T 5' PTP-1B fragment was recovered and ligated
15 with the 500 bp Bgl II/Sal I fragment containing the mutated active site. The ligation was transformed into bacteria (type DH5 α , G) and clones containing the mutated active site sequence identified by sequencing. The GST-PTP-1B Cys²¹⁵Ser mutant was overexpressed, purified and found not to have any phosphatase activity.

20

2. Preparation of Tritiated Bz-NEJJ-CONH₂

This compound can be prepared as outlined in Scheme 1, below, and by following the procedures:

25 Synthesis of N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide (BzN-EJJ-CONH₂)

1.0 g of TentaGel® S RAM resin (RAPP polymer, ~ 0.2 mmol/g) as represented by the shaded bead in Scheme 1, was treated
30 with piperidine (3 mL) in DMF (5 mL) for 30 min. The resin (symbolized by the circular P, containing the remainder of the organic molecule except the amino group) was washed successively with DMF (3 x 10 mL) and CH₂Cl₂ (10 mL) and air dried. A solution of DMF (5 mL), N[∞]-Fmoc-4-[diethylphosphono-(difluoromethyl)]-L-

phenylalanine (350 mg) , where Fmoc is 9-fluorenylmethoxycarbonyl, and O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluranium hexafluorophosphate,(acronym being HATU, 228 mg) was treated with diisopropyl-ethylamine (0.21 mL) and, after 15 min., was added to the resin in 3 mL of DMF. After 1 h, the resin was washed successively with DMF (3x10 mL) and CH₂Cl₂ (10 mL) and air dried. The sequence was repeated two times, first using N[∞]-Fmoc-4-[diethylphosphono-(difluoromethyl)]-L-phenylalanine and then using N-Fmoc-L-glutamic acid gamma-*t*-butyl ester. After the final coupling, the resin bound tripeptide was treated with a mixture of piperidine (3 mL) in DMF (5mL) for 30 min. and was then washed successively with DMF (3x10 mL) and CH₂Cl₂ (10 mL) and air dried.

To a solution of benzoic acid (61 mg) and HATU (190 mg) in DMF (1 mL) was added diisopropylethylamine (0.17 mL) and, after 15 min. the mixture was added to a portion of the resin prepared above (290 mg) in 1 mL DMF. After 90 min. the resin was washed successively with DMF (3 x 10 mL) and CH₂Cl₂ (10 mL) and air dried. The resin was treated with 2 mL of a mixture of TFA: water (9:1) and 0.05 mL of triisopropylsilane (TIPS-H) for 1 h. The resin was filtered off and the filtrate was diluted with water (2 mL) and concentrated *in vacuo* at 35°C. The residue was treated with 2.5 mL of a mixture of TFA:DMS:TMSOTf (5:3:1) and 0.05 mL of TIPS-H, and stirred at 25°C for 15 h. (TFA is trifluoroacetic acid, DMS is dimethyl sulfate, TMSOTf is trimethylsilyl trifluoromethanesulfonate).

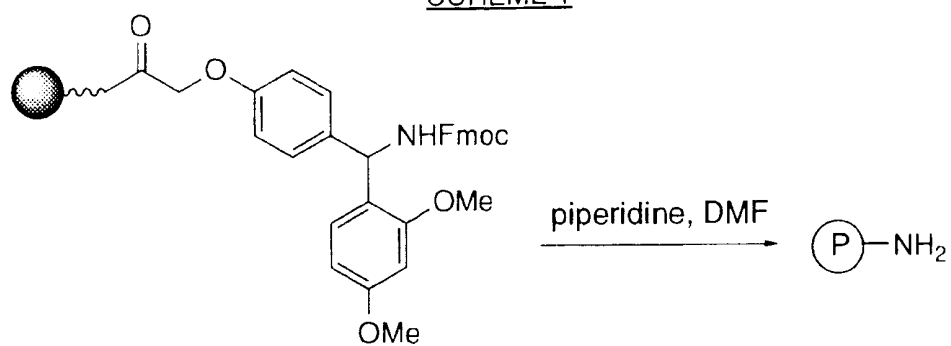
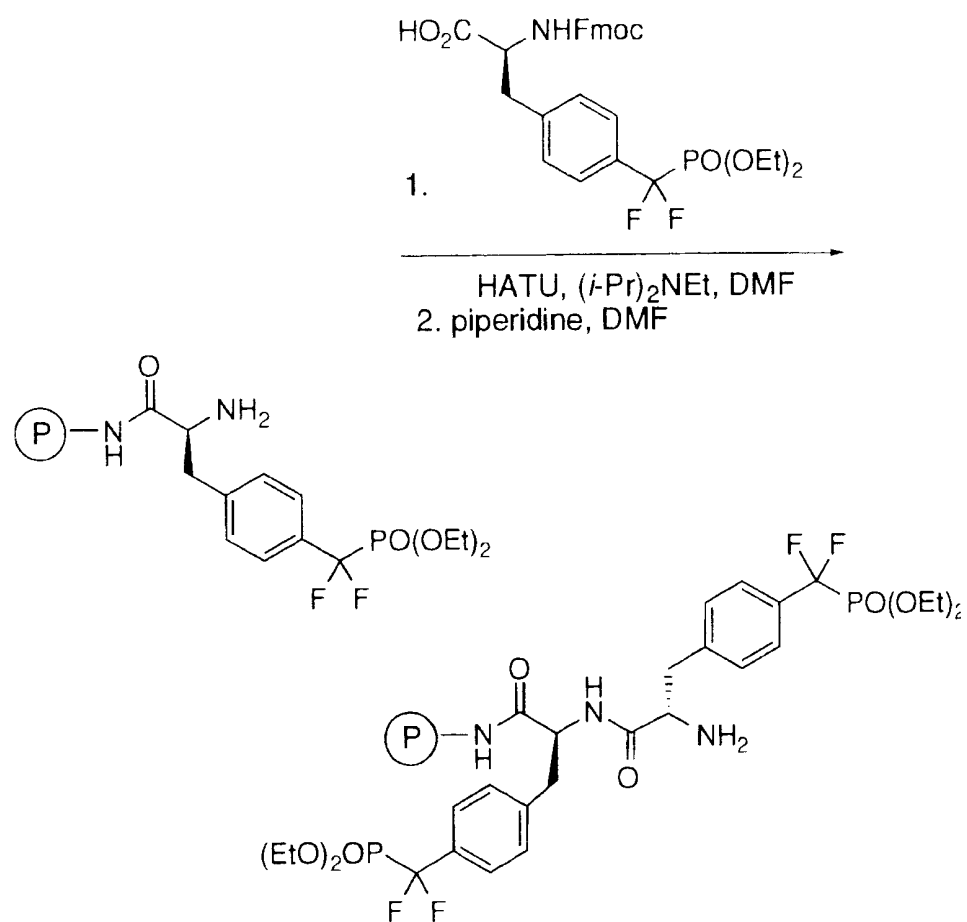
The desired tripeptide, the title compound, was purified by reverse phase HPLC (C18 column, 25 x 100 mm) using a mobile phase gradient from 0.2% TFA in water to 50/50 acetonitrile/0.2% TFA in water over 40 min. and monitoring at 230 nm. The fraction eluting at approximately 14.3 min. was collected, concentrated and lyophilized to yield the title compound as a white foam.

Synthesis of N-(3,5-Ditritio)benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenyl-alanineamide

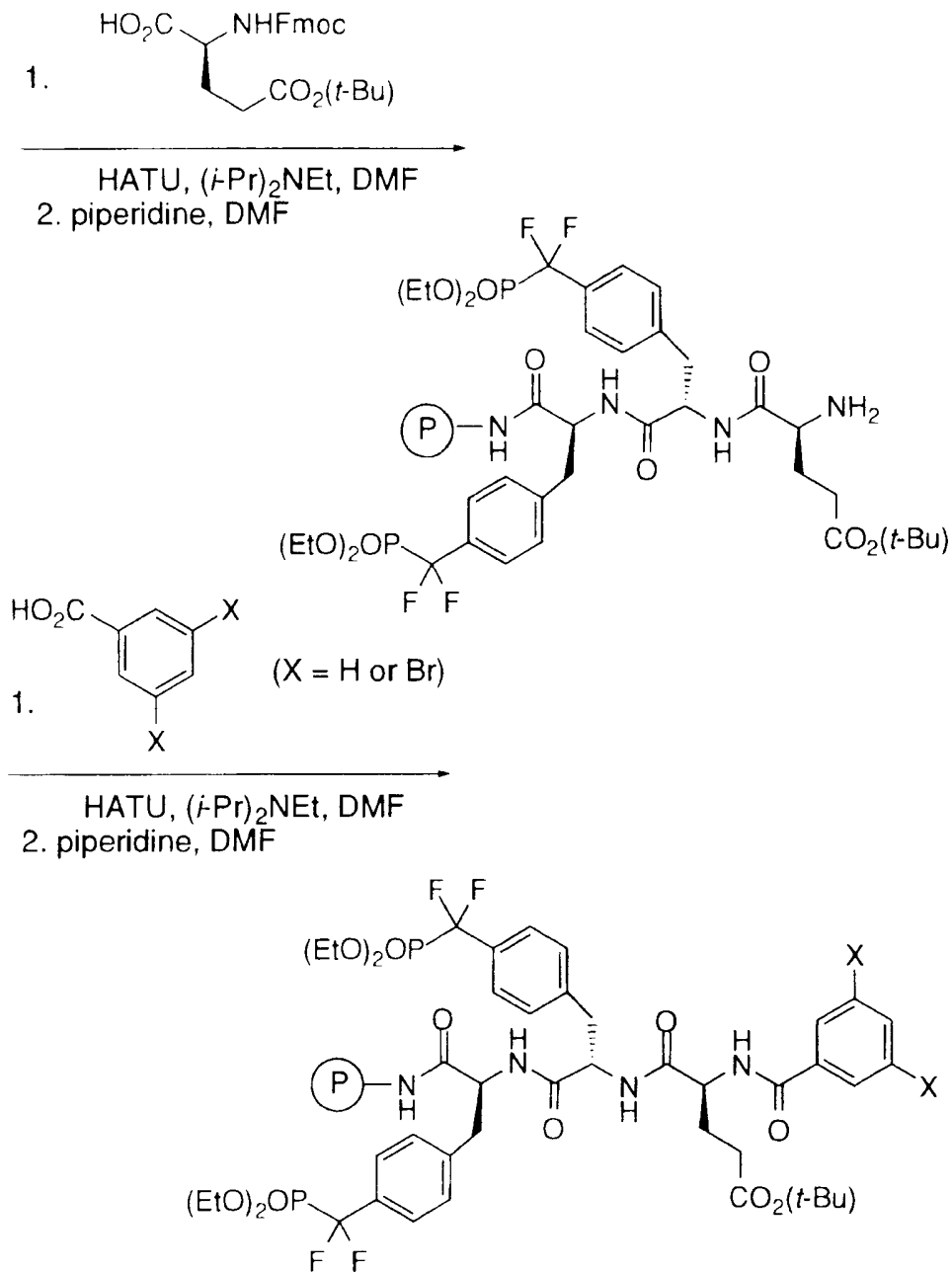
- 5 The above procedure described for the preparation of
BzN-EJJ-CONH₂ was repeated, but substituting 3,5-dibromobenzoic
acid for benzoic acid. After HPLC purification as before, except using
a gradient over 30 min. and collecting the fraction at approximately
18.3 min., the dibromo containing tripeptide was obtained as a white
foam.
- 10 A portion of this material (2 mg) was dissolved in
methanol/triethylamine (0.5 mL, 4/1), 10% Pd-C (2 mg) was added, and the
mixture stirred under an atmosphere of tritium gas for 24 h. The mixture was
filtered through celite, washing with methanol and the filtrate was
concentrated. The title compound was obtained after purification by semi-
- 15 preparative HPLC using a C18 column and an isocratic mobile phase of
acetonitrile/0.2% TFA in water (15:100). The fraction eluting at approximately
5 min. was collected and concentrated *in vacuo*. The title compound was
dissolved in 10 mL of methanol/water (9:1) to provide a 0.1 mg/mL solution of
specific activity 39.4 Ci/mmol.

20

SCHEME 1

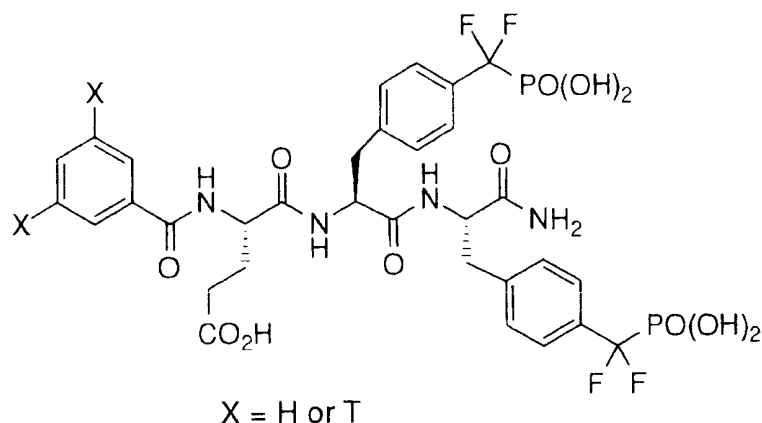
TentaGel[®] S RAM polymer

SCHEME 1 CONT'D



SCHEME 1 CONT'D

1. TFA-H₂O (9:1)
2. TFA-DMS-TMSOTf-TIPSH
3. HPLC purification
4. for X = Br: T₂ (g), 10% Pd-C
MeOH, Et₃N;
HPLC purification



By following the above described procedure for BzN-EJJ-CONH₂, the following other peptide inhibitors were also similarly

- 5 prepared:
N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
10 L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide,
15 L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide, and

L-Isoleucyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide.

4. Phosphatase Assay Protocol

5

Materials:

EDTA - ethylenediaminetetraacetic acid (Sigma)

DMH - N,N'-dimethyl-N,N'-bis(mercaptoacetyl)-
hydrazine (synthesis published in *J. Org. Chem.* 56, pp. 2332-
10 2337,(1991) by R. Singh and G.M. Whitesides and can be substituted
with DTT - dithiothreitol Bistris - 2,2-bis(hydroxymethyl)2,2',2''-
nitrilotriethanol-(Sigma) Triton X-100 - octylphenolpoly(ethylene-
glycolether) 10 (Pierce) Antibody: Anti-glutathione S-transferase
rabbit (H and L) fraction (Molecular Probes) Enzyme: Human
15 recombinant PTP1B, containing amino acids 1-320, (Seq. ID No. 1)
fused to GST enzyme (glutathione S-transferase) purified by affinity
chromatography. Wild type (Seq. ID No. 1) contains active site
cysteine(215), whereas mutant (Seq. ID No. 7) contains active site
serine(215).

20 Tritiated peptide: Bz-NEJJ-CONH₂, Mwt. 808, empirical
formula, C₃₂H₃₂T₂O₁₂P₂F₄

Stock Solutions

25 (10X) Assay Buffer 500 mM Bistris (Sigma), pH 6.2,
MW=209.2
20mM EDTA (GIBCO/BRL)
Store at 4° C.

Prepare fresh daily:

30

Assay Buffer (1X) 50 mM Bistris
(room temp.) 2 mM EDTA
5 mM DMH (MW=208)

Enzyme Dilution

Buffer (keep on ice) 50 mM Bistris
 2 mM EDTA
 5 mM DMH
5 20% Glycerol (Sigma)
 0.01 mg/ml Triton X-100 (Pierce)

Antibody Dilution

Buffer (keep on ice) 50 mM Bistris
10 2 mM EDTA

IC₅₀ Binding Assay Protocol:

Compounds (ligands) which potentially inhibit the
binding of a radioactive ligand to the specific phosphatase are
15 screened in a 96-well plate format as follows:

To each well is added the following solutions @ 25°C in
the following chronological order:

- 20 1. 110 µl of assay buffer.
2. 10 µl. of 50 nM tritiated BzN-EJJ-CONH₂ in assay
buffer (1X) @ 25°C.
3. 10 µl. of testing compound in DMSO at 10 different
concentrations in serial dilution (final DMSO, about 5% v/v) in
duplicate @ 25°C.
- 25 4. 10 µl. of 3.75 µg/ml purified human recombinant
GST-PTP1B in enzyme dilution buffer.
5. The plate is shaken for 2 minutes.
6. 10 µl. of 0.3 µg/ml anti-glutathione S-transferase
(anti-GST) rabbit IgG (Molecular Probes) diluted in antibody dilution
30 buffer @ 25°C.
7. The plate is shaken for 2 minutes.
8. 50 µl. of protein A-PVT SPA beads (Amersham)
@ 25°C.
9. The plate is shaken for 5 minutes. The binding
35 signal is quantified on a Microbeta 96-well plate counter.
10. The non-specific signal is defined as the enzyme-
ligand binding in the absence of anti-GST antibody.

11. 100% binding activity is defined as the enzyme-ligand binding in the presence of anti-GST antibody, but in the absence of the testing ligands with the non-specific binding subtracted.
- 5 12. Percentage of inhibition is calculated accordingly.
13. IC₅₀ value is approximated from the non-linear regression fit with the 4-parameter/multiple sites equation (described in: "Robust Statistics", New York, Wiley, by P.J. Huber (1981) and reported in nM units.
- 10 14. Test ligands (compounds) with larger than 90% inhibition at 10 μ M are defined as actives.

The following Table I illustrates typical assay results of examples of known compounds which competitively inhibit the

15 binding of the binding agent, BzN-EJJ-CONH₂.

TABLE I
GST-PTP1B SPA Binding Assay with Non-Mutated (Cys215) and Mutated enzyme (Ser215)

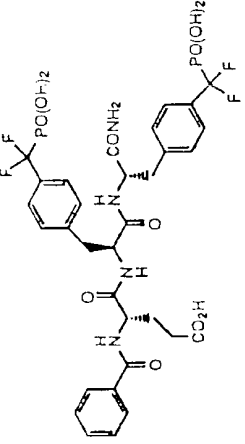
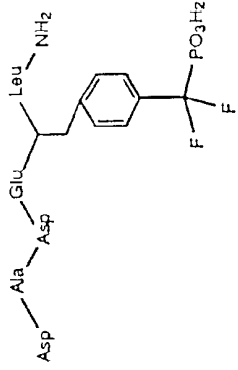
Compound	Structure	Non-Mutated	Mutated
Control:			
Tripeptide(F2PMP)2		14 nM	8 nM
DAD(F2PMP)L hexapeptide (T. Burke et al, Biochem. Biophys. Res. Comm. 204, 129, (1994))		400 nM	100 nM

TABLE I (Cont'd.)

SH-specific binding: Vanadate		2 μ M	>100 μ M
Insulin Receptor Peptide		17 μ M	70 μ M
Potential Oxidizing agents: Hydrogen peroxide	H ₂ O ₂	90% at 83 μ M 4 μ M	0% at 83 μ M >100 μ M
Quinone			
Potential Alkylating agents: Imine		67% at 2 μ M	10% at 2 μ M

TABLE II

Raw Data Counts (dpm)
(duplicates)

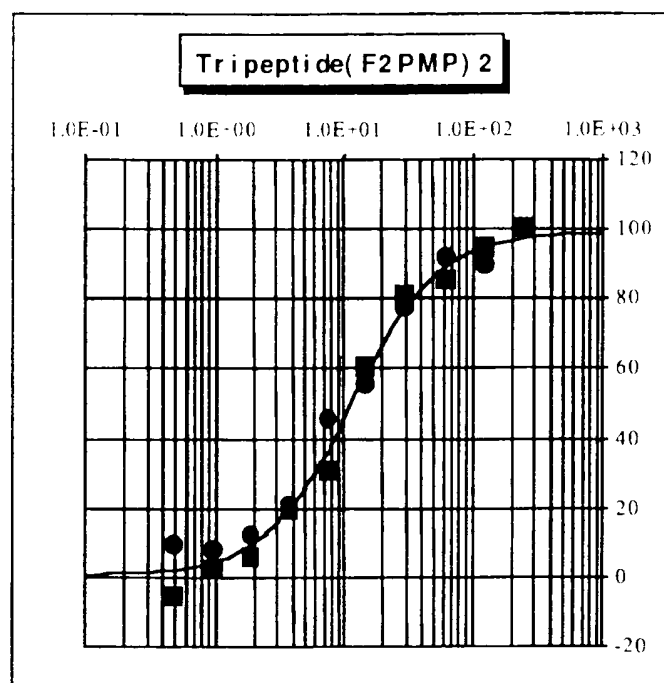
no	antibody			conc. BzN-EJJ-CONH2, nM							
antibody	(+ control)	250	125	62.5	31.25	15.625	7.813	3.906	1.953	0.977	0.488
252	5652	288	873	757	1550	2775	3367	4743	5220	5454	5384
304	6380	273	588	1109	1337	2525	4165	4838	5581	5781	6211

dpm
dpm

TABLE III

no	antibody			conc. Bz-EJJ-CONH2, nM							
antibody	(+ control)	250	125	62.5	31.25	15.625	7.813	3.906	1.953	0.977	0.488
100	5	100	90	92	78	56	45	21	12	8	9
100	-8	100	95	85	81	60	30	19	6	2	-5

% Inh
% Inh



Preparation of Cathepsin K(O2) Mutant (CAT-K Mutant)

Cathepsin K is a prominent cysteine protease in human osteoclasts and is believed to play a key role in osteoclast-mediated bone resorption. Inhibitors of cathepsin K will be useful for the treatment of bone disorders (such as osteoporosis) where excessive bone resorption occurs. Cathepsin K is synthesized as a dormant preproenzyme (Seq. ID No. 4). Both the pre-domain (Met¹-Ala¹⁵) and the prodomain (Leu¹⁶-Arg¹¹⁴) must be removed for full catalytic activity. The mature form of the protease (Ala¹¹⁵-Met³²⁹) contains the active site Cys residue (Cys¹³⁹).

The mature form of cathepsin K is engineered for expression in bacteria and other recombinant systems as a Met Ala¹¹⁵-Met³²⁹ construct by PCR-directed template modification of a clone that is identified. Epitope-tagged variants are also generated: (Met[FLAG]Ala¹¹⁵-Met³²⁹ and Met Ala¹¹⁵-Met³²⁹[FLAG]; where FLAG is the octa-peptide AspTyrLysAspAspAspLys). For the purpose of establishing a binding assay, several other constructs are generated including Met[FLAG]Ala¹¹⁵-[Cys¹³⁹ to Ser¹³⁹]-Met³²⁹ and Met Ala¹¹⁵-[Cys¹³⁹ to Ser¹³⁹]-Met³²⁹[FLAG] (where the active site Cys is mutated to a Ser residue), and Met[FLAG]Ala¹¹⁵-[Cys¹³⁹ to Ala¹³⁹]-Met³²⁹ and Met Ala¹¹⁵-[Cys¹³⁹ to Ala¹³⁹]-Met³²⁹[FLAG] (where the active site Cys is mutated to an Ala residue). In all cases, the resulting re-engineered polypeptides can be used in a binding assay by tethering the mutated enzymes to SPA beads via specific anti-FLAG antibodies that are commercially available (IDI-KODAK). Other epitope tags, GST and other fusions can also be used for this purpose and binding assay formats other than SPA can also be used. Ligands based on the preferred substrate for cathepsin K (e.g. Ac-P2-P₁, Ac-P2-P₁-aldehydes, Ac-P2-P₁-ketones; where P₁ is an amino acid with a hydrophilic side chain, preferably Arg or Lys, and P₂ is an amino acid with a small hydrophobic side chain, preferably Leu, Val or Phe) are suitable in their radiolabeled (tritiated) forms for SPA-based binding assays. Similar binding assays can also be established for other cathepsin family members.

Preparation of Apopain (caspase-3) Mutant

Apopain is the active form of a cysteine protease belonging to the caspase superfamily of ICE/CED-3 like enzymes. It is derived from a catalytically dormant proenzyme that contains both the 17 kDa large subunit (p17) and 12 kDa (p12) small subunit of the catalytically active enzyme within a 32 kDa proenzyme polypeptide (p32). Apopain is a key mediator in the effector mechanism of apoptotic cell death and modulators of the activity of this enzyme, or structurally-related isoforms, will be useful for the therapeutic treatment of diseases where inappropriate apoptosis is prominent, e.g., Alzheimer's disease.

The method used for production of apopain involves folding of active enzyme from its constituent p17 and p12 subunits which are expressed separately in *E. coli*. The apopain p17 subunit (Ser²⁹-Asp¹⁷⁵) and p12 subunit (Ser¹⁷⁶-His²⁷⁷) are engineered for expression as MetSer²⁹-Asp¹⁷⁵ and MetSer¹⁷⁶-His²⁷⁷ constructs, respectively, by PCR-directed template modification. For the purpose of establishing a binding assay, several other constructs are generated, including a MetSer²⁹-[Cys¹⁶³ to Ser¹⁶³]-Asp¹⁷⁵ large subunit and a Met¹-[Cys¹⁶³ to Ser¹⁶³]-His²⁷⁷ proenzyme. In the former case, the active site Cys residue in the large subunit (p17) is replaced with a Ser residue by site-directed mutagenesis. This large subunit is then re-folded with the recombinant p12 subunit to generate the mature form of the enzyme except with the active site Cys mutated to a Ser. In the latter case, the same Cys¹⁶³ to Ser¹⁶³ mutation is made, except that the entire proenzyme is expressed. In both cases, the resulting re-engineered polypeptides can be used in a binding assay by tethering the mutated enzymes to SPA beads via specific antibodies that are generated to recognize apopain (antibodies against the prodomain, the large p17 subunit, the small p12 subunit and the entire p17:p12 active enzyme have been generated). Epitope tags or GST and other fusions could also be used for this purpose and binding assay formats other than SPA can also be used.

Ligands based on the preferred substrate for apopain (variants of AspGluValAsp), such as Ac- AspGluValAsp, Ac- AspGluValAsp-aldehydes, Ac-AspGluValAsp-ketones are suitable

in their radiolabeled forms for SPA-based binding assays. Similar binding assays can also be established for other capsase family members.

5 DESCRIPTION OF THE SEQUENCE LISTINGS

SEQ ID NO. 1 is the top sense DNA strand of Figures 2A and 2B for the PTP1B tyrosine phosphatase enzyme.

- 10 SEQ ID NO. 2 is the amino acid sequence of Figures 2A and 2B for the PTP1B tyrosine phosphatase enzyme.

SEQ ID NO. 3 is the top sense cDNA strand of Figures 3A, 3B and 3C for the Cathepsin K preproenzyme.

15

SEQ ID NO. 4 is the amino acid sequence of Figures 3A, 3B and 3C for the Cathepsin K preproenzyme.

- 20 SEQ ID NO. 5 is the top sense cDNA strand of Figures 4A and 4B for the CPP32 apopain proenzyme.

SEQ ID NO. 6 is the amino acid sequence of Figures 4A and 4B for the CPP32 apopain proenzyme.

- 25 SEQ ID NO. 7 is the cDNA sequence of the human PTP-1B₁₋₃₂₀ Ser mutant.

SEQ ID NO. 8 is the amino acid sequence of the human PTP-1B₁₋₃₂₀ Ser mutant.

30

SEQ ID NO. 9 is the cDNA sequence for apopain C163S mutant.

SEQ ID NO. 10 is the amino acid sequence for the apopain C163S mutant.

35

SEQ ID NO. 11 is the large subunit of the heterodimeric amino acid sequence for the apopain C163S mutant.

SEQ ID NO. 12 is the cDNA sequence for the Cathepsin K C139S mutant.

5 SEQ ID NO. 13 is the cDNA sequence for the Cathepsin K C139A mutant.

SEQ ID NO. 14 is the amino acid sequence for the Cathepsin K C139S mutant.

10

SEQ ID NO. 15 is the amino acid sequence for the Cathepsin K C139A mutant.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: DESMARAIS, SYLVIE
FRIESEN, RICHARD
GRESSEF, MICHAEL
KENNEDY, BRIAN
NICHOLSON, DONALD
RAMACHANDRAN, CHIDAMBARAN
SKOREY, KATHRYN
FORD-HUTCHINSON, ANTHONY

(ii) TITLE OF THE INVENTION: PHOSPHATASE BINDING ASSAY

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: ROBERT J. NORTH - MERCK & CO., INC.
(B) STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
(C) CITY: RAHWAY
(D) STATE: NJ
(E) COUNTRY: USA
(F) ZIP: 07065

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: not known
(B) FILING DATE: 04-NOV-1996
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: NORTH, ROBERT J
(B) REGISTRATION NUMBER: 27,366
(C) REFERENCE/DOCKET NUMBER: 19824 PCT

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 732-594-7262
(B) TELEFAX: 732-594-4720
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 963 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(1) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

ATGGAGATGG AAAAGGAGTT CGAGCAGATG GAGAAGTCCG GGAGCTGGGG GGCCATTTCG    60
CAGGATATGG GAGATGAAGT CAGTGAATTC CCATGTAGAG TGGCCAAAGGT TCCTAAGAAAG    120
AAAAACCGAA ATAGGTACAG AGAGGTCAAT CCTTTGAGG ATAGTGGGAT TAAACTACAT    180
CAAGAAGATA ATGACTATAT CAAGGTCAAT TTGATAAAAA TGAAGAAGG CCAAGGGAAT    240
TACATTCTTA CCGAGGGGTC TTGGCTAAG ACATGCGGTG ACTTTTGGGA GATGGTGTGG    300
GAGCAGAAAA GGAGGGGTGT GGTGATGCTC AACAGAGTGA TGGAGAAAGG TTGGTTAAAA    360
TGGGCACAAT ACTGGCCACA AAAAGAAGAA AAAGAGATGA TTTTGAAGA CACAAATTTG    420
AAATTAACAT TGATCTCTGA AGATATCAAG TCATATTATA CAGTGGGACA GGTAGAATTG    480
GAAAACCTTA CAAGCCAAAG AACTGGAGAG ATTTACATT TCACTATAC CACATGGCCT    540
GACTTTGGAG TGCTGAATG ACCAGCTTCA TTCTTGAAT TTCTTTTCAA AGTCCGAGAG    600
TCAGGGTCAC TCAGCCCGGA CCACGGGTCG GTTGTGGTGC AGTGCAGTGC AGGCATCGGT    660
AGGTCTGGAA CTTCTGTCTT GGTGATAAC TGCTCTCTGC TGATGGACAA GAGGAAAGAG    720
CCTTCTTCGG TTGATATCAA AAAAGTCTG TTAGAAATGA GAAATTTTG CATGGGGTTG    780
ATCCAGACAG CCGACCAGCT GCGCTTCTCT TACCTGCTG TGATGAAGG TGCCAAATTC    840
ATCATGGGGG ACTCTTCGT GAGGATCAG TGAAGGAGG TTCCCAAGG GACCTGGAG    900
CCCCACCGCG AGCATATCC CCACTCTCC CGSCCAACCA AACGAATGCT GGAGCCACAG    960
TGA

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 320 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Glu Met Glu Lys Glu Phe Glu Gln Ile Asp Lys Ser Gly Ser Trp
 1           5           10           15
Ala Ala Ile Tyr Gln Asp Ile Arg His Glu Ala Ser Asp Phe Pro Cys
 20          25          30
Arg Val Ala Lys Leu Pro Lys Asn Lys Asn Arg Asn Arg Tyr Arg Asp
 35          40          45
Val Ser Pro Phe Asp His Ser Arg Ile Lys Leu His Gln Glu Asp Asn
 50          55          60
Asp Tyr Ile Asn Ala Ser Leu Ile Lys Met Glu Glu Ala Gln Arg Ser
 65          70          75          80
Tyr Ile Leu Thr Gln Gly Pro Leu Pro Asn Thr Cys Gly His Phe Trp
 85          90          95

```

Glu Met Val Trp Glu Gln Lys Ser Arg Gly Val Val Met Leu Asn Arg
 100 105 110
 Val Met Glu Lys Gly Ser Leu Lys Cys Ala Gln Tyr Trp Pro Gln Lys
 115 120 125
 Glu Glu Lys Glu Met Ile Phe Glu Asp Thr Asn Leu Lys Leu Thr Leu
 130 135 140
 Ile Ser Glu Asp Ile Lys Ser Tyr Tyr Thr Val Arg Gln Leu Glu Leu
 145 150 155 160
 Glu Asn Leu Thr Thr Gln Glu Thr Arg Glu Ile Leu His Phe His Tyr
 165 170 175
 Thr Thr Trp Pro Asp Phe Gly Val Pro Glu Ser Pro Ala Ser Phe Leu
 180 185 190
 Asn Phe Leu Phe Lys Val Arg Glu Ser Gly Ser Leu Ser Pro Glu His
 195 200 205
 Gly Pro Val Val Val His Cys Ser Ala Gly Ile Gly Arg Ser Gly Thr
 210 215 220
 Phe Cys Leu Ala Asp Thr Cys Leu Leu Leu Met Asp Lys Arg Lys Asp
 225 230 235 240
 Pro Ser Ser Val Asp Ile Lys Lys Val Leu Leu Glu Met Arg Lys Phe
 245 250 255
 Arg Met Gly Leu Ile Gln Thr Ala Asp Gln Leu Arg Phe Ser Tyr Leu
 260 265 270
 Ala Val Ile Glu Gly Ala Lys Phe Ile Met Gly Asp Ser Ser Val Gln
 275 280 285
 Asp Gln Trp Lys Glu Leu Ser His Glu Asp Leu Glu Pro Pro Pro Glu
 290 295 300
 His Ile Pro Pro Pro Pro Arg Pro Pro Lys Arg Ile Leu Glu Pro His
 305 310 315 320

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1669 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAACAAGCA CTGGATTCCA TATCCCACTG CCAAAACCGC ATGGTTCAGA TTATCGGTAT 60
 TGCAGCTTTC ATCATAATAC ACAGCTTTGC TGCCGAAACG AAGCCAGACA ACAGATTTCC 120
 ATCAGCAGGA TGTGGGGGCT CAAGCTTCTG CTGCTACCTG TGGTGAGCTT TGCTCTGTAC 180
 CCTGAGGAGA TACTGGACAC CCACTGGGAG CTATGGAAGA AGACCCACAG GAAGCAATAT 240
 AACAACAAGG TGGATGAAAT CTCTGGGGCT TTAATTGGG AAAAAACCT GAAGTATATT 300
 TCCATCCATA ACCTTGAGGC TTCTTTTGGT GTCCATACAT ATGAAGTGGC TATGAACCAC 360
 CTGGGGGACA TGACCACTGA AGAGGTGGTT CAGAAGATGA CTGGAATCAA AGTACCCCTG 420
 TCTCATTCCT GCAGTAATGA CAGCTTTTAT ATCCGAGAAT GGAAGGTAG AGGCCAGAC 480
 TCTGTGACT ATCGAAAGAA AGGATATGTT ACTCTGTCA AAAATCAGG TCAATGTGGT 540
 TCTGTGGGG CTTTGTAGCT TGTGGGTGCC CTGGAGGGCC AACTCAAGAA GAAACTGGC 600
 AAACCTTTAA ATCTGAGTCC CCAGAACCTA GTGGATTGTG TGTCTGAGAA TGATGGCTGT 660

```

GGAGGGGGT ATATGACCA TGGCTTGGAA TATGTGAGA AGAAGGGGG TATTGACTGT 720
GAAGATGAGT AAGTATATGT GGGACAGGAA GAGATTGTA TGTATAAGG AACAGGGAAG 760
GGAGTAAGT GGAGAGGGA GAGAGAGATC GGGAGGGGA ATGAGAAAGT TGTGAAGAGG 800
GGAGTGGG GAGTGGGAGT TGTCTGTGTG GGTATTGATG CAAGCGTGAG CTGCTTGGAG 840
TTTTACAGCA AAGGTGTGTA TTATGATGAA AGGTGCAATA GGGATAATGT GAACCATGGG 880
GTTTTGGGAG TGGGATATGG AATGGAGAAG GGAACAAGC ATGGGATAAT TAAAAACAGC 920
TGGGGAGAAA AATGGGGAAA GAAAGGATAT ATCTGTATGG CTGGAAATAA GAACAAGGGG 960
TGTGGGATTT GGAAGCTGGG GAGCTTGGCC AAGATGTGAG TGGAGCGAGG GAAATGGATG 1000
CTGCTCTTTC ATTTCTTCCA GATGGGTGCA GTGTAAAGAT GCACTTTGGA AGGGAGTTGG 1040
TGTGCTATTT TTGAAGCAGA TGTGGTGATA CTGAGATTGT CTGTTGAGTT TGGCCATTGG 1080
TTTGTGCTTC AATGATGCTT TCGTACTTTG CTTCTCTCCA GCGATGACCT TTTTCACTGT 1120
GGCATGAGG AATTTCTCTG ACAGCTGTGT ACTTTAGGC TAAGAGATGT GACTACAGCG 1160
TGGCGCTGAT TGTGTTTCTC CAGGGCTGAT GCTGTACAGG TACAGGCTGG AGATTTTCAC 1200
ATAGGTTAGA TTCTCATTTA GGGGACTAGT TAGCTTTAAG CAGCCTAGAG GACTAGGGTA 1240
ATCTGACTTC TCACTTCTTA AGTTGCTTTC TATATCTCA AGSTAGAAAT GTCTATGTTT 1280
TCTACTGCAA TTATAAATC TATTCATAAG TCTTTGGTAC AAGTTTACAT GATAAAAAGA 1320
AATGTGATTT GTTTTCCCTT CTTTGCACCT TTGAAATAAA GTATTTATC 1360

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
 1           5           10          15
Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Lys Thr
          20          25          30
His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu
          35          40          45
Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala
          50          55          60
Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
          65          70          75          80
Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro
          85          90          95
Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr Ile Pro Glu Trp Glu
          100         105         110
Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys Lys Gly Tyr Val Thr
          115         120         125
Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ser
          130         135         140

```

Val Gly Ala Leu Gln Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
 145 150 155 160
 Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly
 165 170 175
 Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Lys Asn
 180 185 190
 Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu Glu
 195 200 205
 Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr
 210 215 220
 Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala
 225 230 235 240
 Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala Ser Leu Thr Ser Phe
 245 250 255
 Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Ser Cys Asn Ser Asp
 260 265 270
 Asn Leu Asn His Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly
 275 280 285
 Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly Glu Asn Trp Gly Asn
 290 295 300
 Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Ala Cys Gly Ile
 305 310 315 320
 Ala Asn Leu Ala Ser Phe Pro Lys Met
 325

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1001 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTGCAGGAAT TCGGCACGAG GGGTGTATT GTGAGGCGGT TGTAGAACTT AATAAAGGTA 60
 TCCATGGAGA ACACTGAAAA CTCAGTGGAT TCAAAATCCA TTAATAATTT GGAACCAAAG 120
 ATCATACATG GAAGCGAATC AATGGACTCT GGAATATCCC TGGACAACAG TTATAAAATG 180
 GATTATCCTG AGATGGGTTT ATGTATAATA ATTAATAATA AGAATTTTCA TAAGAGCACT 240
 GGAATGACAT CTCGGTCTTG TACAGATGTC GATGCAGCAA ACCTCAGGGA AACATTCAGA 300
 AACTTGAAAT ATGAAGTCAG GAATAAAAAAT GATCTTACAC GTGAAGAAAT TGTGGAATTG 360
 ATGGCGTGATG TTTCTAAAGA AGATCACAGC AAAAGGAGCA GTTTTGTGTG TGTGCTTCTG 420
 AGCCATGGTG AAGAAGGAAT AATTTTGGGA ACAAATGGAC CTGTTGACCT GAAAAAATA 480
 ACAAACTTTT TCAGAGGGGA TCGTTGTAGA AGTCTAACTG GAAAAACCAA ACTTTTCATT 540
 ATTCAGGCGT GCGGTGTAC AGAACTGGAC TGTGSCATTG AGACAGACAG TGGTGTGTGAT 600
 GATGACATGG CGTGTGATAA AATACCAGTG GAGGCGGACT TCTGTATGC ATACTCCACA 660
 GCACCTGGTT ATTATTCTTG GCGAAATTCA AAGGATGGCT CCTGTTTAT CCAGTGGCTT 720
 TGTGCCATG TGAAACAGTA TCCCGACAAG CTTGAATTTA TGCACATTCT TACCCGGGTT 780

```

AATGGAAGAG TGGAACAGA ATTGAGTGG TTTTCTTTT AGGTACTTT TCATGCAAG 540
AAGAGATT CATGTATTGT CCGCATGTC ACAAAGAGAG TGTATTTTTA TCACTAAGA 560
AATGTTTGGT TGSTGGTTT TTTAGTTTG TATGCAAGT GAGAAGATG TATATTTGGT 580
ACTGATTTT CCGTCATTT TGACCTACT TCATGCTGA G 1061

```

(2) INFORMATION FOR SEQ ID NO:6:

1. SEQUENCE CHARACTERISTICS:

- A. LENGTH: 277 amino acids
- B. TYPE: amino acid
- C. STRANDEDNESS: single
- D. TOPOLOGY: linear

11. MOLECULE TYPE: peptide

(xx) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Glu Asn Thr Glu Asn Ser Val Asp Ser Lys Ser Ile Lys Asn Leu
1      5      10      15
Glu Pro Lys Ile Ile His Gly Ser Glu Ser Met Asp Ser Gly Ile Ser
20      25      30
Leu Asp Asn Ser Tyr Lys Met Asp Tyr Pro Glu Met Gly Leu Cys Ile
35      40      45
Ile Ile Asn Asn Lys Asn Phe His Lys Ser Thr Gly Met Thr Ser Arg
50      55      60
Ser Gly Thr Asp Val Asp Ala Ala Asn Leu Arg Glu Thr Phe Arg Asn
65      70      75      80
Leu Lys Tyr Glu Val Arg Asn Lys Asn Asp Leu Thr Arg Glu Glu Ile
85      90      95
Val Glu Leu Met Arg Asp Val Ser Lys Glu Asp His Ser Lys Arg Ser
100     105     110
Ser Phe Val Cys Val Leu Leu Ser His Gly Glu Glu Gly Ile Ile Phe
115     120     125
Gly Thr Asn Gly Pro Val Asp Leu Lys Lys Ile Thr Asn Phe Phe Arg
130     135     140
Gly Asp Arg Cys Arg Ser Leu Thr Gly Lys Pro Lys Leu Phe Ile Ile
145     150     155     160
Gln Ala Cys Arg Gly Thr Glu Leu Asp Cys Gly Ile Glu Thr Asp Ser
165     170     175
Gly Val Asp Asp Asp Met Ala Cys His Lys Ile Pro Val Glu Ala Asp
180     185     190
Phe Leu Tyr Ala Tyr Ser Thr Ala Pro Gly Tyr Tyr Ser Trp Arg Asn
195     200     205
Ser Lys Asp Gly Ser Trp Phe Ile Gln Ser Leu Cys Ala Met Leu Lys
210     215     220
Gln Tyr Ala Asp Lys Leu Glu Phe Met His Ile Leu Thr Arg Val Asn
225     230     235     240
Arg Lys Val Ala Thr Glu Phe Glu Ser Phe Ser Phe Asp Ala Thr Phe
245     250     255
His Ala Lys Lys Gln Ile Pro Cys Ile Val Ser Met Leu Thr Lys Glu
260     265     270
Leu Tyr Phe Tyr His
275

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 963 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

ATGGAGATGG AAAAGGAGTT CGAGCAGATC GACAAGTCCG GGAGCTGGGC GGCCATTTAC      60
CAGGATATCC GACATGAAGC CAGTGAATTG CCATGTAGAG TGGCCAAGCT TCCTAAGAAC      120
AAAAACCGAA ATAGGTACAG AGACGTGAGT CCCTTTGACC ATAGTCGGAT TAAACTACAT      180
CAAGAAGATA ATGACTATAT CAACGCTAGT TTGATAAAAA TGGAAGAAGC CCAAAGGAGT      240
TACATTCTTA CCCAGGCCCC TTGCGCTAAC ACATGCGGTC ACTTTTGGGA GATGGTGTGG      300
GAGCAGAAAA GCAAGGGTGT CGTCATGCTC AACAGAGTGA TGGAGAAAGG TTCGTTAAAA      360
TGCGCACAAT ACTGGCCACA AAAAGAAGAA AAAGAGATGA TCTTTGAAGA CACAAATTTG      420
AAATTAACAT TGATCTCTGA AGATATCAAG TCATATTATA CAGTGCGACA GCTAGAATTG      480
GAAAACTTTA CAACCCCAAG AACTCGAGAG ATCTTACATT TCCACTATAC CACATGGCCT      540
GACTTTGAGG TCCCTGAATC ACCAGCCTCA TTCTTGAAGT TTCTTTTCAA AGTCCGAGAG      600
TCAGSGTCAG TCAGCCCGGA GCACGGGCCC GTTGTGGTGC ACAGCAGTGC AGGCATCGGC      660
AGGTCTGAAA CTTCTCTCTT GGGTATACCC TGGCTCTGTC TGATGGACAA GAGGAAAGAC      720
CCTTCTTTCG TTGATATCAA GAAAGTGCTG TTAGAAATGA GGAAGTTTCG GATGGGGTTG      780
ATCCAGACAG CCGACCAAGT GCGCTTCTCC TACCTGGCTG TGATCGAAGG TGCCAAATTC      840
ATCATGGGGG ACTTTTCCGT GCAGGATCAG TGGAAGGAGC TTCCCCACGA GGACCTGGAG      900
CCCCCAAGCG AGCATATCCC CCCACCTCCC CGGCCACCCA AACGAATCCT GGAGCCACAC      960
TGA

```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 322 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Glu Met Glu Lys Glu Phe Glu Gln Ile Asp Lys Ser Gly Ser Trp
 1           5           10           15
Ala Ala Ile Tyr Gln Asp Ile Arg His Glu Ala Ser Asp Phe Pro Cys
 20           25           30
Arg Val Ala Lys Leu Pro Lys Asn Lys Asn Arg Asn Arg Tyr Arg Asp
 35           40           45

```

Val Ser Pro Phe Asp His Ser Arg Ile Lys Leu His Gln Glu Asp Asn
 50 55 60
 Asp Tyr Ile Asn Ala Ser Leu Ile Lys Met Glu Glu Ala Gln Arg Ser
 65 70 75 80
 Tyr Ile Leu Thr Gln Gly Pro Leu Pro Asn Thr Cys Gly His Phe Trp
 85 90 95
 Glu Met Val Trp Glu Gln Lys Ser Arg Gly Val Val Met Leu Asn Arg
 100 105 110
 Val Met Glu Lys Gly Ser Leu Lys Cys Ala Gln Tyr Trp Pro Gln Lys
 115 120 125
 Glu Glu Lys Glu Met Ile Phe Glu Asp Thr Asn Leu Lys Leu Thr Leu
 130 135 140
 Ile Ser Glu Asp Ile Lys Ser Tyr Tyr Thr Val Arg Gln Leu Glu Leu
 145 150 155 160
 Glu Asn Leu Thr Thr Gln Glu Thr Arg Glu Ile Leu His Phe His Tyr
 165 170 175
 Thr Thr Trp Pro Asp Phe Gly Val Pro Glu Ser Pro Ala Ser Phe Leu
 180 185 190
 Asn Phe Leu Phe Lys Val Arg Glu Ser Gly Ser Leu Ser Pro Glu His
 195 200 205
 Gly Pro Val Val Val His Ser Ser Ala Gly Ile Gly Thr Cys Gly Arg
 210 215 220
 Ser Gly Thr Phe Cys Leu Ala Asp Thr Cys Leu Leu Met Asp Lys
 225 230 235 240
 Arg Lys Asp Pro Ser Ser Val Asp Ile Lys Lys Val Leu Leu Glu Met
 245 250 255
 Arg Lys Phe Arg Met Gly Leu Ile Gln Thr Ala Asp Gln Leu Arg Phe
 260 265 270
 Ser Tyr Leu Ala Val Ile Glu Gly Ala Lys Phe Ile Met Gly Asp Ser
 275 280 285
 Ser Val Gln Asp Gln Trp Lys Glu Leu Ser His Glu Asp Leu Glu Pro
 290 295 300
 Pro Pro Glu His Ile Pro Pro Pro Arg Pro Pro Lys Arg Ile Leu
 305 310 315 320
 Glu Pro

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1001 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGCAGGAAT TCGGCACGAG GGSTGCTATT GTGAGGCGGT TGTAGAAGTT AATAAAGGTA 60
 TCCATCGAGA ACACTGAAAA CTCAGTGGAT TCAAAATCCA TTAAAAATTT GGAACCAAAG 120
 ATCATACATG GAAGCGAATC AATGGACTCT GGAATATCCC TGGACAACAG TTATAAAATG 180
 GATTATCCTG AGATGGSTTT ATGTATAATA ATTAATAATA AGAATTTTCA TAAGAGCACT 240
 GGAATGACAT CTCGGTCTGG TACAGATGTC GATGCAGCAA ACCTCAGGGA AACATTGAGA 300
 AACTTGAAAT ATGAAGTCAG GAATAAAAAT GATCTTACAC GTGAAGAAAT TGTGGAATTG 360


```

ATGCGTGATG TTTCTAAAGA AGATCAGAGC AAAAGGAGCA GTTTTSTTTG TGTGTTTGT 420
AGCCATGGTG AAGAAGGAAT AATTTTGGG ACAAATGGAC CTGTTGACCT GAAAAAATA 480
ACAAACTTTT TCAGAGGGGA TCGTTGTAGA AATCTAACTG GAAAACCCAA ACTTTTCATT 540
ATTCAGGCCT CCGTGTGTAC AGAACTGGAC TGTGSCATTG AGACAGACAG TGGTGTGTAT 600
GATGACATGG CGTGTATAA AATACCAGTG GAGGCGGACT TCTGTATGC ATACTCCACA 660
GCACCTGGTT ATTATCTTG GCGAAATCA AAGGATGGCT CCTGGTTTAT CCAGTGGCTT 720
TGTGCCATGC TGAACAGTA TGCGGACAAG CTTGAATTTA TGCACATTCT TACCCGGGTT 780
AACCAGAAAG TGGCAACAGA ATTTGAGTCC TTTTCTTTG ACGCTACTTT TCATGCAAAG 840
AAACAGATTC CATGTATTGT TTCCATGCTC ACAAAGAAG TCTATTTTTA TCACTAAAGA 900
AATGCTTGGT TGGTGGTTTT TTTAGTTTG TATGCAAGT GAGAAGATGG TATATTGGT 960
ACTGTATTTC CCTCTCATTT TGACCTACTC TCATGCTGCA G 1001

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Glu Asn Thr Glu Asn Ser Val Asp Ser Lys Ser Ile Lys Asn Leu
 1          5          10          15
Glu Pro Lys Ile Ile His Gly Ser Glu Ser Met Asp Ser Gly Ile Ser
 20          25          30
Leu Asp Asn Ser Tyr Lys Met Asp Tyr Pro Glu Met Gly Leu Cys Ile
 35          40          45
Ile Ile Asn Asn Lys Asn Phe His Lys Ser Thr Gly Met Thr Ser Arg
 50          55          60
Ser Gly Thr Asp Val Asp Ala Ala Asn Leu Arg Glu Thr Phe Arg Asn
 65          70          75          80
Leu Lys Tyr Glu Val Arg Asn Lys Asn Asp Leu Thr Arg Glu Glu Ile
 85          90          95
Val Glu Leu Met Arg Asp Val Ser Lys Glu Asp His Ser Lys Arg Ser
100          105          110
Ser Phe Val Cys Val Leu Leu Ser His Gly Glu Glu Gly Ile Ile Phe
115          120          125
Gly Thr Asn Gly Pro Val Asp Leu Lys Lys Ile Thr Asn Phe Phe Arg
130          135          140
Gly Asp Arg Cys Arg Ser Leu Thr Gly Lys Pro Lys Leu Phe Ile Ile
145          150          155          160
Gln Ala Ser Arg Gly Thr Glu Leu Asp Cys Gly Ile Glu Thr Asp Ser
165          170          175
Gly Val Asp Asp Asp Met Ala Cys His Lys Ile Pro Val Glu Ala Asp
180          185          190
Phe Leu Tyr Ala Tyr Ser Thr Ala Pro Gly Tyr Tyr Ser Trp Arg Asn
195          200          205
Ser Lys Asp Gly Ser Trp Phe Ile Gln Ser Leu Cys Ala Met Leu Lys
210          215          220

```

```

Gln Tyr Ala Asp Lys Leu Glu Phe Met His Ile Leu Thr Arg Val Asn
235      236      238      240
Arg Lys Val Ala Thr Gln Phe Glu Ser Ile Ser Phe Asp Ala Thr Phe
245      250      255
His Ala Lys Lys Gln Ile Pro Cys Ile Val Ser Met Leu Thr Lys Glu
260      265      270
Leu Tyr Phe Tyr His
275

```

2. INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

Met Glu Asn Thr Glu Asn Ser Val Asp Ser Lys Ser Ile Lys Asn Leu
 1      5      10      15
Glu Pro Lys Ile Ile His Gly Ser Glu Ser Met Asp Ser Gly Ile Ser
      20      25      30
Leu Asp Asn Ser Tyr Lys Met Asp Tyr Pro Glu Met Gly Leu Cys Ile
      35      40      45
Ile Ile Asn Asn Lys Asn Phe His Lys Ser Thr Gly Met Thr Ser Arg
      50      55      60
Ser Gly Thr Asp Val Asp Ala Ala Asn Leu Arg Glu Thr Phe Arg Asn
      65      70      75      80
Leu Lys Tyr Glu Val Arg Asn Lys Asn Asp Leu Thr Arg Glu Glu Ile
      85      90      95
Val Glu Leu Met Arg Asp Val Ser Lys Glu Asp His Ser Lys Arg Ser
      100      105      110
Ser Phe Val Cys Val Leu Leu Ser His Gly Glu Glu Gly Ile Ile Phe
      115      120      125
Gly Thr Asn Gly Pro Val Asp Leu Lys Lys Ile Thr Asn Phe Phe Arg
      130      135      140
Gly Asp Arg Cys Arg Ser Leu Thr Gly Lys Pro Lys Leu Phe Ile Ile
      145      150      155      160
Gln Ala Ser Arg Gly Thr Glu Leu Asp Cys Gly Ile Glu Thr Asp Ser
      165      170      175
Gly Val Asp Asp Asp Met Ala Cys His Lys Ile Pro Val Glu Ala Asp
      180      185      190
Phe Leu Tyr Ala Tyr Ser Thr Ala Pro Gly Tyr Tyr Ser Trp Arg Asn
      195      200      205
Ser Lys Asp Gly Ser Trp Phe Ile Gln Ser Leu Cys Ala Met Leu Lys
      210      215      220
Gln Tyr Ala Asp Lys Leu Glu Phe Met His Ile Leu Thr Arg Val Asn
      225      230      235      240
Arg Lys Val Ala Thr Glu Phe Glu Ser Phe Ser Phe Asp Ala Thr Phe
      245      250      255
His Ala Lys Lys Gln Ile Pro Cys Ile Val Ser Met Leu Thr Lys Glu
      260      265      270
Leu Tyr Phe Tyr His
      275

```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

ATGTGGGGGC TCAAGTTTCT GCTGCTACCT GTGGTGAGCT TTGCTCTGTA CCCTGAGGAG      60
ATACTGGACA CCCACTGGGA GCTATGGAAG AAGACCCACA GGAAGCAATA TAACAACAAG      120
GTGGATGAAA TCTCTGGGCG TTAAATTTGG GAAAAAACC TGAAGTATAT TTCCATCCAT      180
AACCTTGAGG CTTCTCTTGG TGCCATACA TATGAAGTGG CTATGAACCA CCTGGGGGAG      240
ATGACCAATG AAGAGTTGGT TCAGAAGATG ACTGGACTCA AAGTACCCCT GTCTCATTCG      300
CGCAGTAATG ACACCTTTTA TATCCCAAGT TGGGAAGGTA GAGCCCGAGA CTCTCTGGAG      360
TATCGAAAGA AAGGATATCT TACTCCTCTC AAAAATCAGG CTCAGTGTGG TTCCTCTTGG      420
GCTTTTAACT CTGTGGGTGC CCTGGAGGGC CAACTCAAGA AGAAAACTGG CAAACTCTTA      480
AATCTGAGTC CCCAGAACCT AGTGGATTGT GTGTCTGAGA ATGATGGCTG TGGAGGGGGC      540
TACATGACCA ATCCCTTCCA ATATGTGCAG AAGAACCAGG GTATTGACTC TGAAGATGCC      600
TACCCATATG TGGGACAGGA AGAGAGTTGT ATGTACAACC CAACAGGCAA GGCAGCTAAA      660
TGCAGAGGGT ACAGAGAGAT CCCCGAGGGG AATGAGAAAG CCCTGAAGAG GGCAGTGGGC      720
CGAGTGGGAC CTGTCTCTGT GGGTATTGAT GCAAGGCTGA CCTCCTTCCA GTTTTACAGC      780
AAAGTGTTCT ATTATGATGA AAGCTGCAAT AGCGATAATC TGAACCATGC GGTTTTGGCA      840
GTGGGATATG GAATCCAGAA GGGAAACAAG CACTGGATAA TTA AAAACAG CTGGGGAGAA      900
AACTGGGGAA ACAAAGGATA TATCCTCATG GCTCGAAATA AGAACAACGC CTGTGGCATT      960
GCCAACCTGG CCAGCTTCCC CAAGATCTGA

```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

ATGTGGGGGC TCAAGTTTCT GCTGCTACCT GTGGTGAGCT TTGCTCTGTA CCCTGAGGAG      60
ATACTGGACA CCCACTGGGA GCTATGGAAG AAGACCCACA GGAAGCAATA TAACAACAAG      120
GTGGATGAAA TCTCTGGGCG TTAAATTTGG GAAAAAACC TGAAGTATAT TTCCATCCAT      180

```

```

AACCTTGAGG CTTCTCTTTC CTTCTATATA TATGAAGTGG CTATGAAGCA CTTGGGGGAG 140
ATGAGCAAGT AAGAGGTGGT TGAAGAAATG ACTCAATTA AAGTACGCTT CTCTCATTCG 150
AGCAGTAATG ACAGCTTTTA TATCCAGAGT TGGGAAGSTA GAGCCGAGAG CTCTGTGAGC 160
TATCGAAAGA AAGGATATGT TACTCCTGTC AAAAATCAGG CTCAGTGTGG TTCCGCTTGG 170
GCTTTTAGGT CTGTGGGTGG CTTGGAGGGG CAACTCAAGA AGAAAAGTGG CAAATCTTTA 180
AATCTGAGTC CCGAGAACCT AGTGGATTGT GTCTCTGAGA ATGATGGGTG TGGAGGGGGC 190
TACATGAGCA ATCTCTTCCA ATATGTGAG AAGAAACGGG GTATTGATTC TGAAGATGCC 200
TACCATATG TGGAACAGGA AGAGAGTTGT ATGTACAACC CAACAGGCAA GGCAGCTAAA 210
TGCAGAGGGT ACAGAGAGAT CCGGAGGGG AATGAGAAAG CCGTGAAGAG GGCAGTGGCC 220
CGAGTGGGAC CTGTCTGTGT GGCATTGAT GCAAGCCTGA CCTCCTTCCA GTTTTACAGC 230
AAAGGTGTGT ATTATGATGA AAGGTGCAAT AGCGATAATC TGAACCATGC GGTTTTGGCA 240
GTGGGATATG GAATCCAGAA GGGAAACAAG CACTGGATAA TTAATAATAG CTGGGAGAGAA 250
AACTGGGGAA ACAAGGATA TATCTCATG GCTCGAAATA AGAACAATGC CTGTGGCATT 260
GCCAACCTGG CCGAGTTCCC CAAGATGTGA 270

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
1      5      10      15
Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Lys Thr
20      25      30
His Arg Lys Lys Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu
35      40      45
Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala
50      55      60
Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
65      70      75      80
Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro
85      90      95
Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr Ile Pro Glu Trp Glu
100     105     110
Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys Lys Gly Tyr Val Thr
115     120     125
Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Ser Trp Ala Phe Ser Ser
130     135     140
Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
145     150     155     160
Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly
165     170     175

```

Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Lys Asn
 120 185 190
 Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu Glu
 195 200 205
 Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr
 210 215 220
 Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala
 225 230 235 240
 Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala Ser Leu Thr Ser Phe
 245 250 255
 Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Ser Cys Asn Ser Asp
 260 265 270
 Asn Leu Asn His Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly
 275 280 285
 Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly Glu Asn Trp Gly Asn
 290 295 300
 Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Asn Ala Cys Gly Ile
 305 310 315 320
 Ala Asn Leu Ala Ser Phe Pro Lys Met
 325

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
 1 5 10 15
 Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Thr
 20 25 30
 His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu
 35 40 45
 Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala
 50 55 60
 Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
 65 70 75 80
 Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro
 85 90 95
 Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr Ile Pro Glu Trp Glu
 100 105 110
 Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys Lys Gly Tyr Val Thr
 115 120 125
 Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Ala Trp Ala Phe Ser Ser
 130 135 140
 Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
 145 150 155 160
 Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly
 165 170 175
 Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Lys Asn
 180 185 190
 Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu Glu
 195 200 205

```

Ser  Tyr  Met  Tyr  Asn  Pro  Thr  Gly  Lys  Ala  Ala  Lys  Cys  Arg  Gly  Tyr
210                      215                      220
Arg  Gln  Ile  Pro  Gln  Gly  Asn  Gln  Lys  Ala  Leu  Lys  Arg  Ala  Val  Ala
225                      230                      235                      240
Arg  Val  Gly  Pro  Val  Ser  Val  Ala  Ile  Asp  Ala  Ser  Leu  Thr  Ser  Phe
245                      250                      255
Gln  Phe  Tyr  Ser  Lys  Gly  Val  Tyr  Tyr  Asp  Glu  Ser  Cys  Asn  Ser  Asp
260                      265                      270
Asn  Leu  Asn  His  Ala  Val  Leu  Ala  Val  Gly  Tyr  Gly  Ile  Gln  Lys  Gly
275                      280                      285
Asn  Lys  His  Trp  Ile  Ile  Lys  Asn  Ser  Trp  Gly  Glu  Asn  Trp  Gly  Asn
290                      295                      300
Lys  Gly  Tyr  Ile  Leu  Met  Ala  Arg  Asn  Lys  Asn  Asn  Ala  Cys  Gly  Ile
305                      310                      315                      320
Ala  Asn  Leu  Ala  Ser  Phe  Pro  Lys  Met
325

```

WHAT IS CLAIMED:

1. A process for determining the binding ability of a ligand to a cysteine-containing wild-type enzyme comprising the steps of:
 - (a) contacting a complex with the ligand, the complex comprising a mutant form of the wild-type enzyme, in which cysteine, at the active site, is replaced with serine, in the presence of a known binding agent for the mutant enzyme, wherein the binding agent is capable of binding with the mutant enzyme to produce a measurable signal.
2. The process of Claim 1 further comprising, the step of contacting the complex with the binding agent, in the absence of the ligand, to produce a first measurable signal.
3. The process of Claim 1 wherein the signal is a colorimetric, photometric, spectrophotometric or radioactive signal.
4. The process of Claim 3 wherein the signal is a beta radiation-induced scintillation.
5. The process of Claim 1 wherein the known binding agent is an inhibitor for the wild-type enzyme and contains a radionuclide to induce scintillation upon contact with the mutant enzyme.
6. The process of Claim 1 wherein the complex further comprises a solid support, a scintillation agent, and a fused enzyme linking construct.
7. The process of Claim 6 wherein the complex is further comprised of:
 - (a) a fluopolymer bead containing a scintillation agent and Protein A, which is attached via Protein A to;

- (b) an anti-GST antibody, which is further attached to the GST end of;
- (c) a fused enzyme linking construct comprised of GST enzyme fused with the mutant enzyme.

5

8. The process of Claim 1 wherein the wild-type enzyme is selected from the group consisting of proteases, phosphatases, lipases, hydrolases and kinases.

10

9. The process of Claim 8 wherein the wild-type enzyme is selected from the group consisting of tyrosine phosphatases and cysteine proteases.

15

10. The process of Claim 9 wherein the tyrosine phosphatase is selected from the group consisting of PTP1B, LCA, LAR, DLAR and DPTP.

20

11. The process of Claim 11 wherein the tyrosine phosphatase is PTP1B which contains serine in place of cysteine at position 215.

25

12. The process of Claim 11 wherein the PTP1B phosphatase is present in a truncated form comprising amino acids 1-320 and containing the active binding site.

30

13. The process of Claim 9 wherein the cysteine protease is a Cathepsin or capsase.

14. The process of Claim 13 wherein the cathepsin is selected from the group consisting of Cathepsin B, Cathepsin G, Cathepsin J, Cathepsin K(O2), Cathepsin L, Cathepsin M and Cathepsin S.

35

15. The process of Claim 14 wherein the cathepsin is Cathepsin K(O2).

16. The process of Claim 11 wherein the capsase is selected from the group consisting of : capsase-1(ICE), capsase-2 (ICH-1), capsase-3 (CPP32, human apopain, Yama), capsase-4(ICE_{rel}-11, TX, ICH-2), capsase-5(ICE_{rel}-111, TY), capsase-6(Mch2),
- 5 capsase-7(Mch3, ICE-LAP3, CMH-1), capsase-8(FLICE, MACH, Mch5), capsase-9 (ICE-LAP6, Mch6) and capsase-10(Mch4).

17. The process of Claim 16 wherein the capsase is human apopain CPP32 .

10

18. The process of Claim 11 wherein the tyrosine phosphatase is PTP1B and the binding agent is a peptide selected from the group consisting of:
- 15 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide (BzN-EJJ-CONH₂), where E is glutamic acid and J is 4-phosphono(difluoro-methyl)-L-phenylalanyl;
- N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 20 N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 25 L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide; and
- 30 L-Isoleucinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide.

19. The process of Claim 18 wherein the peptide is in tritiated form.

35

20. The process of Claim 18 wherein the peptide is tritiated N-(3,5-Ditritio)benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide, being tritiated Bz-NEJJ-CONH₂, wherein E as used herein is glutamic acid and J, as used
 5 herein, is the (F₂Pmp) moiety, (4-phosphono-(difluoromethyl)phenylalanyl).

21. A process for determining the binding ability of a ligand to a cysteine-containing wild-type tyrosine phosphatase comprising the steps of:

- 10 (a) contacting a complex with the ligand, the complex comprising a mutant form of the wild-type enzyme, the mutant enzyme being PTP1B, containing the same amino acid sequence 1-320 as the wild type enzyme, except at position 215, in which cysteine is
 15 replaced with serine in the mutant enzyme, in the presence of a known radioligand binding agent for the mutant enzyme, wherein the binding agent is capable of binding with the mutant enzyme to produce a measurable beta radiation-induced
 20 scintillation signal.

22. The process of Claim 21 further comprising before step (a), the step of contacting the complex with the binding agent in the absence of the ligand to produce a first measurable beta radiation-induced scintillation
 25 signal.

23. The process of Claim 21 wherein the binding agent is a peptide selected from the group consisting of:
 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-
 30 phosphono(difluoromethyl)]-L-phenylalanineamide (BzN-EJJ-CONH₂), where E is glutamic acid and J is 4-phosphono(difluoro-methyl)]-L-phenylalanyl;
 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
 N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-
 35 phosphono(difluoromethyl)]-L-phenylalanine amide;
 L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;

- L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono-
(difluoromethyl)]-L-phenylalanine amide;
L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono-
(difluoromethyl)]-L-phenylalanine amide;
- 5 L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono-
(difluoromethyl)]-L-phenylalanine amide; and
L-Isoleucinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono-
(difluoromethyl)]-L-phenylalanine amide.
- 10 24. The process of Claim 23 wherein the peptide is in tritiated
or 125 I iodinated form.
25. The process of Claim 24 wherein the peptide is tritiated N-
(3,5-Ditritio)benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-
15 phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide, being
tritiated Bz-NEJJ-CONH₂, wherein E as used herein is glutamic acid and J, as
used herein, is the (F₂Pmp) moiety, (4-phosphono-
(difluoromethyl)phenylalanyl).
- 20 26. A complex comprised of:
- (a) a mutant form of a wild-type enzyme, in which
cysteine, necessary for activity in the active site, is
replaced with serine and is attached to:
- (b) a solid support.
- 25 27. The complex of Claim 26 further comprising: a binding
agent for the mutant enzyme, wherein the binding agent is capable of binding
with the mutant enzyme to produce a measurable signal.

28. A peptide selected from the group consisting of:

- N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanineamide (BzN-EJJ-CONH₂), where E is glutamic acid and J is 4-phosphono(difluoro-methyl)]-L-phenylalanyl;
- 5 N-Benzoyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- N-Acetyl-L-glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 10 L-Glutamyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Lysinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- L-Serinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide;
- 15 L-Prolinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide; and
- L-Isoleucinyl-[4-phosphono(difluoromethyl)]-L-phenylalanyl-[4-phosphono(difluoromethyl)]-L-phenylalanine amide, for use as a binding agent for a mutant enzyme.

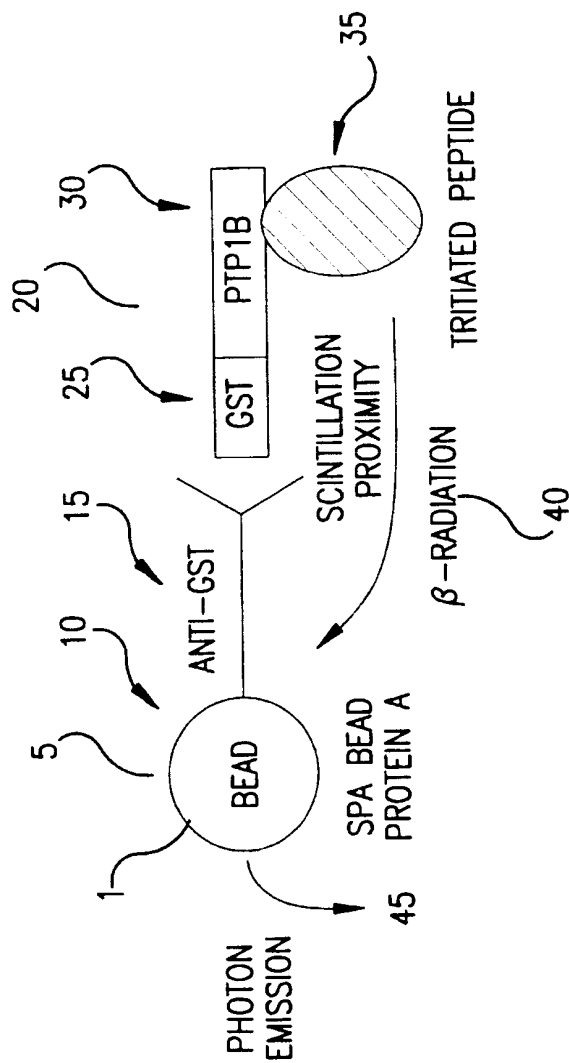


FIG.1

278

1	ATGGAGATGGAAAAAGAGTTCCGAGCAGATCGACAAGTCCGGGAGCTGGGGGGCCATTTAC	30
1	TACTTCTACCTTTTCTCAAGCTCGTCTAGCTGTTCAAGGCCCTCGACCCGCCGGTAAATG	20
1	MetGluMetGluLysGluPheGluGlnIleAspLysSerGlySerTrpAlaAlaIleTyr	
61	CAGGATATCCGACATGAAGCCAGTGACTTCCCATGTAGAGTGGCCAAGCTTCTTAAGAAC	120
61	GTCCTATAGGCTGTACTTCGGTCACTGAAGGGTACATCTCAGCGGTTGGAAGGATTCTTG	40
21	GlnAspIleArgHisGluAlaSerAspPheProCysArgValAlaLysLeuProLysAsn	
121	AAAAACCGAAATAGGTACAGAGACGTCACTCCCTTTGACCATAGTCGGATTAACTACAT	180
121	TTTTTGGCTTTATCCATGTCTCTGCAGTCAGGGAACTGGTATCAGCCTAATTTGATGTA	60
41	LysAsnArgAsnArgTyrArgAspValSerProPheAspHisSerArgIleLysLeuHis	
181	CAAGAAGATAATGACTATATCAACGCTAGTTTGATAAAAATGGAAGAAGCCCAAAGGAGT	240
61	GTTCTTCTATTACTGATATAGTTGGGATCAAACCTATTTTTACCTTCTTCGGGTTTCTCTCA	80
61	GlnGluAspAsnAspTyrIleAsnAlaSerLeuIleLysMetGluGluAlaGlnArgSer	
241	TACATTCTTACCCAGGGGCCCTTTGCCTAACACATGGGGTCACTTTTGGGAGATGGTGTGG	300
241	ATGTAAGAATGGGTCCCGGGAAACGGATTGTGTACGCCAGTGAAAACCTCTACCACACC	160
81	TyrIleLeuThrGlnGlyProLeuProAsnThrCysGlyHisPheTrpGluMetValTrp	
301	GAGCAGAAAAGCAGGGGTGTCGTCAATGCTCAACAGAGTGATGGAGAAAGGTTGTTAAAA	360
301	CTCGTCTTTTCTGTCCTCACAGCAGTACGAGTTGTCTCACTACCTCTTTCCAAGCAATTTT	120
161	GluGlnLysSerArgGlyValValMetLeuAsnArgValMetGluLysGlySerLeuLys	
361	TGCGCACAATACTGGCCACAAAAAGAAGAAAAAGAGATGATCTTTGAAGACACAAATTTG	420
361	ACCGGTGTATGACCGGTGTTTTTCTCTTTTCTCTACTAGAACTTCTGTGTTTAAAC	140
121	CysAlaGlnTyrTrpProGlnLysGluGluLysGluMetIlePheGluAspThrAsnLeu	
421	AAATTAACATTGATCTCTGAAGATATCAAGTCATATTATACAGTGGACAGCTAGAATTG	480
421	TTTAATTGTAAGTAGAGACTTCTATAGTTCAGTATAATATGTCACGCTGTGGATCTTAAC	160
141	LysLeuThrLeuIleSerGluAspIleLysSerTyrTyrThrValArgGlnLeuGluLeu	
481	GAAAACTTACAACCAAGAACTCGAGAGATCTTACATTTCCTACTATACCACATGGCCT	540
481	CTTTTGGAAATGTTGGGTTCTTTGAGCTCTCTAGAATGTAAGGTGATATGGTGTACGGGA	180
161	GluAsnLeuThrThrGlnGluThrArgGluIleLeuHisPheHisTyrThrThrTrpPro	

FIG.2A

3/8

541 GACTTGGAGTCCCTGAATCACCAGCCTCATTCTTGAACCTTCTTTTCAAAGTCCGAGAG 600
-----+-----+-----+-----+-----+-----+
181 CTGAAACCTCAGGGACTTAGTGGTCGGAGTAAGAACTTGAAAGAAAAGTTTCAGGCTCTC 700
AspPheGlyValProGluSerProAlaSerPheLeuAsnPheLeuPheLysValArgGlu
601 TCAGGGTCACTCAGCCCGGAGCACGGGCCCGTTGTGGTGAAGTGCAGTGCAGGCATCGGC 600
-----+-----+-----+-----+-----+-----+
201 AGTCCCAGTGAGTCGGGCCTCGTGCCCGGGCAACACCACGTGACGTACGTCCGTAGCCG 200
SerGlySerLeuSerProGluHisGlyProValValValHisCysSerAlaGlyIleGly
661 AGGTCTGGAACCTTCTGTCTGGCTGATACCTGCCTCCTGCTGATGGACAAGAGGAAAAGAC 720
-----+-----+-----+-----+-----+-----+
221 TCCAGACCTTGGAGACAGACCGACTATGGACGGAGGACGACTACCTGTTCTCCTTTCTG 240
ArgSerGlyThrPheCysLeuAlaAspThrCysLeuLeuLeuMetAspLysArgLysAsp
721 CCTCTTCCGTGATATCAAGAAAGTGCTGTTAGAAATGAGGAAGTTTCGGATGGGGTTG 780
-----+-----+-----+-----+-----+-----+
241 GGAAGAAGGCAACTATAGTTCTTTCACGACAATCTTTACTCCTTCAAAGCCTACCCCAAC 260
ProSerSerValAspIleLysLysValLeuLeuGluMetArgLysPheArgMetGlyLeu
781 ATCCAGACAGCCGACCAGCTGCGCTTCTCCTACCTGGCTGTGATCGAAGGTGCCAAATTC 840
-----+-----+-----+-----+-----+-----+
261 TAGTCTGTGCGCTGGTCGACGCGAAGAGGATGGACCGACACTAGCTTCCACGGTTTAAG
IleGlnThrAlaAspGlnLeuArgPheSerTyrLeuAlaValIleGluGlyAlaLysPhe
841 ATCATGGGGGACTCTTCCGTGCAGGATCAGTGGAAGGAGCTTTCCACGAGGACCTGGAG 900
-----+-----+-----+-----+-----+-----+
TAGTACCCCTGAGAAGGCACGTCCTAGTCACCTTCTCGAAAGGGTGCTCCTGGACCTC
IleMetGlyAspSerSerValGlnAspGlnTrpLysGluLeuSerHisGluAspLeuGlu
901 CCCGCACCCGAGCATATCCCCCACCTCCCCGGCCACCCAAACGAATCCTGGAGCCACACTGA 960
-----+-----+-----+-----+-----+-----+
301 GGGGGTGGGCTCGTATAGGGGGTGGAGGGGCCGGTGGGTTTGCTTAGGACCTCGGTGTGACT 320
ProProProGluHisIleProProProProArgProProLysArgIleLeuGluProHisEnd

FIG.2B

1/3

1 GAAACAAGCAGCTGGATTCCATATCGGCACTGCCAAAACCGCATGGTTGAGATTATCGCTAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 61 CTTTGTTCGTGACCTAAGGTATAGGGTGACGGTTTGGGGTAGCAAGTCTAAATAGCGATA
 TGCAGCTTTTCATCATAATACACACCTTTGCTGCGGAAACGAAGCCAGACAACAGATTTCC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 121 ACGTCGAAAGTAGTATATGTGTGGAACGACGGCTTTGCTTCGGTCTGTGTCTAAAGG
 ATCAGTAGGATGTGGGGGCTCAAGGTTCTGCTGCTACCTGTGGTGAGTTTGGTCTGTAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 181 TAGTGTCTACACCCCCGAGTTCCAAGACGACGATGGACACCACTCGAAACGAGACATG
 MetTrpGlyLeuLysValLeuLeuLeuProValValSerPheAlaLeuTyr
 CCTGAGGAGATACTGGACACCCACTGGGAGCTATGGAAGAAGACCCACAGGAAGCAATAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 241 GGACTTCCTCTATGACCTGTGGGTGACCCCTCGATACCTTCTTCTGGGTGTCTTGGTTATA
 ProGluGluIleLeuAspThrHisTrpGluLeuTrpLysLysThrHisArgLysGlnTyr
 AACAAACAAGGTGGATGAAATCTCTCGGCGTTTAATTTGGGAAAAAAACCTGAAGTATATT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 301 TTGTTGTTCACCTACTTTAGACAGCCGCAAATTAAACCTTTTTTTGGACTTCATATAA
 AsnAsnLysValAspGluIleSerArgArgLeuIleTrpGluLysAsnLeuLysTyrIle
 TCCATCCATAACCTTGAGGCTTCTCTTGGTGTCCATACATATGAACTGGCTATGAACCAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 361 AGGTAGGTATTGGAACCTCGAAGAGAACACAGGTATGTATACTTGACCGATACTTGGTG
 SerIleHisAsnLeuGluAlaSerLeuGlyValHisThrTyrGluLeuAlaMetAsnHis
 CTGGGGGACATGACCAGTGAAGAGGTGGTTCAGAAGATGACTGGAAGTCAAAGTACCCCTG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 421 GACCCCTGTACTGGTCACTTCTCCACCAAGTCTTCTACTGACCTGAGTTTCATGGGGAC
 LeuGlyAspMetThrSerGluGluValValGlnLysMetThrGlyLeuLysValProLeu
 TCTCATTCGGCAGTAATGACACCCCTTTATATCCCAGAATGGGAAGGTAGAGCCCCAGAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 481 AGAGTAAGGGCGTCATTACTGTGGGAAATATAGGGTCTTACCCTTCCATCTCGGGGTCTG
 SerHisSerArgSerAsnAspThrLeuTyrIleProGluTrpGluGlyArgAlaProAsp
 TCTGTGCACTATCGAAAGAAAGGATATGTTACTCCTGTCAAAAATCAGGGTCAGTGTGGT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 541 AGACAGCTGATAGCTTTCTTTCTATACAATGAGGACAGTTTTTAGTCCCAGTCACACCA
 SerValAspTyrArgLysLysGlyTyrValThrProValLysAsnGlnGlyGlnCysGly

FIG.3A

5/8

TCCTGTTGGGCTTTAGCTCTGTGGGTGCCCTGGAGGGCCAACTCAAGAAGAAAACCTGGC
541 -----+-----+-----+-----+-----+-----+-----+ 600
AGGACAACCCGAAAATCGAGACACCCACGGGACCTCCCGGTIGAGTTCTTCTTTGACCG
SerCysTrpAlaPheSerSerValGlyAlaLeuGluGlyGlnLeuLysLysLysThrGly
139

AAACTCTTAAATCTGAGTCCCCAGAACCTAGTGGATTGTGTGTCTGAGAAATGATGCTGT
601 -----+-----+-----+-----+-----+-----+-----+ 660
TTTGAGAATTTAGACTCAGGGGTCTTGGATCACCTAACACACAGACTCTTACTACCGACA
LysLeuLeuAsnLeuSerProGlnAsnLeuValAspCysValSerGluAsnAspGlyCys

GGAGGGGGCTACATGACCAATGCCTTCCAATATGTGCAGAAGAACCAGGGGTATTGACTCT
661 -----+-----+-----+-----+-----+-----+-----+ 720
CCTCCCCGATGTACTGGTTACGGAAGGTTATACACGTCTTCTTGGCCCCATAACTGAGA
GlyGlyGlyTyrMetThrAsnAlaPheGlnTyrValGlnLysAsnArgGlyIleAspSer

GAAATGCCTACCCATATGTGGGACAGGAAGAGAGTTGTATGTACAACCCAACAGGCAAG
721 -----+-----+-----+-----+-----+-----+-----+ 780
CTTCTACGGATGGGTATACACCTGTCTCTCTCAACATACATGTGGGTGTCCGTTCC
GluAspAlaTyrProTyrValGlyGlnGluGluSerCysMetTyrAsnProThrGlyLys

GCAGCTAAATGCAGAGGGTACAGAGAGATCCCGAGGGGAATCAGAAAGCCCTGAAGAGG
781 -----+-----+-----+-----+-----+-----+-----+ 840
CGTCGATTTACGTCTCCCATGTCTCTCTAGGGGCTCCCTTACTCTTTCGGGACTTCTCC
AlaAlaLysCysArgGlyTyrArgGluIleProGluGlyAsnGluLysAlaLeuLysArg

GCAGTGGCCCGAGTGGGACCTGTCTCTGTGGCCATTGATGCAAGCCTGACCTCCTTCCAG
841 -----+-----+-----+-----+-----+-----+-----+ 900
CGTCACCGGGCTCACCCCTGGACAGAGACACCGGTAACCTACGTTCTGACTGGAGGAAGGTC
AlaValAlaArgValGlyProValSerValAlaIleAspAlaSerLeuThrSerPheGln

TTTTACAGCAAAGGTGTGTATTATGATGAAAGCTGCAATAGCGATAATCTGAACCATGCG
901 -----+-----+-----+-----+-----+-----+-----+ 960
AAAATGTCGTTTCCACACATAATACTACTTTCGACGTTATCGCTATTAGACTTGGTACGC
PheTyrSerLysGlyValTyrTyrAspGluSerCysAsnSerAspAsnLeuAsnHisAla

GTTTTGCCAGTGGGATATGGAATCCAGAAGGGAACAAGCACTGGATAATTAAAAACAGC
961 -----+-----+-----+-----+-----+-----+-----+ 1020
CAAAACCGTCAACCTATACCTTAGGTCTTCCCTTTGTCGTGACCIATTAATTITGTGCG
ValLeuAlaValGlyTyrGlyIleGlnLysGlyAsnLysHisTrpIleIleLysAsnSer

TGCGGAGAAAACCTGGGGAAACAAAGGATATATCTCATGGCTCGAAATAAGAACAAACGCC
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
ACCCCTCTTTTGACCCCTTTGTTTCCTATATAGGAGTACCGAGCTTATTCTTGTTCGGG
TrpGlyGluAsnTrpGlyAsnLysGlyTyrIleLeuMetAlaArgAsnLysAsnAsnAla

FIG.3B

6/8

1081 TGTGGGATGGCCAAAGCTGGCCAGGTTGGCCAAAGATGGAGTCCAGCCAGGCCAAATCTATCG 1140
 -----+-----+-----+-----+-----+-----+-----+
 ACAACGGTAAAGGTTGGACCGGTGGAAAGCGGTTCTACACTGAGGTGGGTTGGTTTAGGTAG
 CysGlyIleAlaAsnLeuAlaSerPheProLyMetEnd

 1141 CTGCTGTTCCATTCTTCCACCATGGTGGAGTGTAAAGATGCACTTTGTAAGGGAGTTGG 1200
 -----+-----+-----+-----+-----+-----+-----+
 GACGAGAAAGTAAAGCAAGGTGTATCCAGGTCAATTGCTACGTGAAAGCTTCCCTCAACG

 1201 TGTGCTATTTTGAAGCAGATGTGGTGATACTGAGATTGTCTGTTCAGTTTCCCATTTGG 1260
 -----+-----+-----+-----+-----+-----+-----+
 ACAAGATAAAAAGTTCTGTCTACAGTACTATGAGTCTAACAGACAAGTCAAAGGGGTAAAG

 1261 TTTGTGCTTCAATGATGCTTCTTCTTGTCTCTCTACCCATCAGCTTTTCTACTGTI 1320
 -----+-----+-----+-----+-----+-----+-----+
 AAACACGAAAGTTTACTACGAAGGATGAAAGCAAGAGAGGTGGGTACTGGAAAAAGTGACA

 1321 GGCCATCAGGACITTCCTTGACAGTGTGTGACCTTAGGCTAAGAGATGTGACTACAGCC 1380
 -----+-----+-----+-----+-----+-----+-----+
 CCGGTAGTCTGAAAGGGAGTGTGACACATGAGAATCGGATCTCTATACTGATGTGGG

 1381 TGCCCCCTGACTGTGTTGTGTCAGGGCTGATGCTGTACAGGTACAGGCTGGAGATTTTCAC 1440
 -----+-----+-----+-----+-----+-----+-----+
 ACGGGGACTGACACAACAGGGTCCCGACTACGACATGTCCATGTCCGACCTCTAAAAGTG

 1441 ATAGGTTAGATTCTCATTCACGGGACTAGTTAGCTTTAAGCACCTAGAGGACTAGGGTA 1500
 -----+-----+-----+-----+-----+-----+-----+
 TATCCAATCTAAGAGTAAGTGCCCTGATCAATCGAAATTGGTGGGATCTCCTGATCCCAT

 1501 ATCTGACTTCTCACTTCTTAAGTTCTCTTCTATATCCTCAAGGTAGAAATGTCTATGTT 1560
 -----+-----+-----+-----+-----+-----+-----+
 TAGACTGAAGAGTGAAGGATTCAAGGGAAGATATAGGAGTTCATCTTTACAGATACAAA

 1561 TCTACTCCAATTCTAAATCTATTCTATAAGTCTTTGGTACAAGTTTACATGATAAAAAGA 1620
 -----+-----+-----+-----+-----+-----+-----+
 AGATGAGGTAAAGTATCTAGATAAGTATTCAGAAACATGTTCAAATGACTATTTTCTI

 1621 AATGTGATTTGTCTTCCCTTCTTTGCACTTTTGAATAAAGTATTTATC 1669
 -----+-----+-----+-----+-----+-----+-----+
 TTACACTAAACAGAAGGGAAGAAAGGTGAAAACCTTATTTCTAATAATAG

FIG.3C

7/8

1	CTGCAGGAATTCCGGCAGGAGGGGTGCTATTCGAGGCGGTTGTAGAAGTAAATAAAGGTA -----+-----+-----+-----+-----+-----+-----+ CACGTCCTTAAGCCGTGCTCCCTACGATAAACTCCGCCAACAATCTCAATTATTTCCAT	60
61	TCCATGGAGAACACAGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACTAAAAG -----+-----+-----+-----+-----+-----+-----+ AGGTACCTCTTGTGACTTTTGTAGTCACCTAAGTTTAAAGTAATTTTAAACCTTGGTTTC MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLys	120
121	ATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATG -----+-----+-----+-----+-----+-----+-----+ TAGTATGTACCTTCGCTTAGTTACCTGAGACCTTATAGGACCTGTTGCAATATTTTAC IleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnSerTyrLysMet	180
181	GATTATCCTGAGATCGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACT -----+-----+-----+-----+-----+-----+-----+ CTAATAGGACTCTACCCAAATACATAATTAATTATTTATTCTTAAAAGTATCTCGTGA AspTyrProGluMetGlyLeuCysIleIleIleAsnAsnLysAsnPheHisLysSerThr	240
241	GGAATGACATCTCGGTCTGGTACAGATGTGGATGCAGCAAACTCAGGGAAACATTCAGA -----+-----+-----+-----+-----+-----+-----+ CCCTACTGTAGAGCCAGACCATGTCTACAGCTACGTGGTTTGCAGTCCCTTTGTAAGTCT GlyMetThrSerArgSerGlyThrAspValAspAlaAlaAsnLeuArgGluThrPheArg	300
301	AACCTGAAATATGAAGTCAGGAATAAAAAAGATCTTAACGTTGAAGAAATGTGGAATTG -----+-----+-----+-----+-----+-----+-----+ TTGAACTTTATACTTCAGTCCCTATTTTTTACTAGAAAGGCACTTCCTTAAACACCTTAAC AsnLeuLysTyrGluValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeu	360
361	ATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAAACAGTTTTGTTGTGTGCTTCTG -----+-----+-----+-----+-----+-----+-----+ TACGCACTACAAAGATTTCTTCTAGTGTGTTTTCTCTGTCAAAACAAACACACGAAGAC MetArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLeu	420
421	AGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAATA -----+-----+-----+-----+-----+-----+-----+ TCGGTACCACTTCCTCTATTAATAAACCTTGTTTACCTGGACAACAGGACTTTTATTAT SerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeuLysLysIle	480
481	ACAAACTTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATT -----+-----+-----+-----+-----+-----+-----+ TGTGTTGAAAAAGTCTCCGCTAGCAACATCTTCAGATTGACCTTTTGGCTTGGAAAAGTAA ThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGlyLysProLysLeuPheIle	540

FIG.4A

8-6

541 ATTCAGGCCCTGCCGTGGTACAGAACTGGACTGTGGCAATGAGACAGACAGTGGTGTTGAT
 -----+-----+-----+-----+-----+-----+-----+
 TAAGTCCGACGAGGACCATGTCTTGAACCGACACGTAACCTGTCTGTACCCACAACCTA
 IleuIleAlaCysArgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyValAsp
 163

601 GATGACATGGCGTGTATATAAAATACGAGTGGAGGCCGACTTCTTGTATGCATACCTCCACA
 -----+-----+-----+-----+-----+-----+-----+
 CTACTGTACCGCACATTAATTTATGGTTCACCTCCCGGTGAACAACATACGTATGAGGTGT
 AspAspMetAlaCysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThr

661 GCAGCTGGTTATTTATCTTGGCGAAATTCAAAGGATGGGCTCTGGTTCATCCAGTCGCTT
 -----+-----+-----+-----+-----+-----+-----+
 CGTGGACCAATAATAAGAACCGCTTTAAGTTTCCTACCGAGGACCAAGTAGGTCAGCGAA
 AlaProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLeu

721 TGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTATGCACATTCCTTACCGGGTT
 -----+-----+-----+-----+-----+-----+-----+
 ACACGGTAGCACTTTGTCTACCGGTGTTTGAACCTAAATACGTGTAAGAAAGGCCCCAA
 CysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrArgVal

781 AACCGAAAGGTGGCAACAGAATTTGAGTGCCTTTTCCTTTACCGCTACTTTTCATGCAAAG
 -----+-----+-----+-----+-----+-----+-----+
 TTGGCTTTCCACCGTTGTCTTAAACTCAGGAAAAAGAAAGTCGGATGAAAAGTACGTTTC
 AsnArgLysValAlaThrGluPheGluSerPheSerPheAspAlaThrPheHisAlaLys

841 AAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACCTATTTTTATCACTAAAGA
 -----+-----+-----+-----+-----+-----+-----+
 TTTGTCTAAGGTACATAACAAAGGTACGAGTGTTTCTTGAGATAAAAATAGTGATTTCT
 LysGlnIleProCysIleValSerMetLeuThrLysGluLeuTyrPheTyrHisEnd

901 AATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGT
 -----+-----+-----+-----+-----+-----+-----+
 TTACCAACCAACCAACCAAAAAAATCAAACATACGGTTCACCTCTTCTACCATATAAACCA

961 ACTGTATTTCCCTCTCATTTTGACCTACTCTCATGCTGCAG
 -----+-----+-----+-----+-----+-----+-----+ 1001
 TGACATAAAGGGGAGAGTAAACTGGATGAGAGTACGACGTC

FIG.4B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 97/00825

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/34 C12N11/00 C07K7/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 20039 A (CALSBURG A S) 27 July 1995 see the whole document ---	1,21,26
A	EP 0 571 939 A (BOEHRINGER MANNHEIM GMBH) 1 December 1993 see the whole document ---	1,21
A	EP 0 313 244 A (ONCOGEN) 26 April 1989 see the whole document ---	1,21
A	T YOSHIMURA ET AL.: "Substitution of S-(beta-aminoethyl)-cysteine for active-site lysine of thermostable aspartate aminotransferase." JOURNAL OF BIOCHEMISTRY, vol. 108, no. 5, November 1990, pages 699-700, XP002057018 see the whole document ---	1,21,26
-/--		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 February 1998

Date of mailing of the international search report

06/03/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5816 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx: 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Moreno, C

INTERNATIONAL SEARCH REPORT

In International Application No.

PCT/CA 97/00825

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
A	<p>T. R. BURKE ET AL: "Potent inhibition of insulin receptor dephosphorylation by a hexamer peptide containing the phosphotyrosyl mimetic F2Pmp." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 204, no. 1, 14 October 1994, ORLANDO, FL US, pages 129-134, XP002057019 cited in the application see the whole document -----</p>	28

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 97/00825

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9520039 A	27-07-95	AU 679861 B AU 2651395 A CA 2174525 A EP 0723585 A JP 9504438 T	10-07-97 08-08-95 27-07-95 31-07-96 06-05-97
EP 0571939 A	01-12-93	DE 4217474 A JP 2603422 B JP 6046893 A US 5434054 A	02-12-93 23-04-97 22-02-94 18-07-95
EP 0313244 A	26-04-89	AT 123148 T DE 3853866 D DE 3853866 T ES 2073402 T JP 1237454 A US 5646002 A	15-06-95 29-06-95 21-09-95 16-08-95 21-09-89 08-07-97

